

Result No.	Query			Description			
	Score	Match	Length	DB	ID		
C	1	952.8	61.6	1164	1	Q03159	Complete gene sequ
	2	946.2	61.2	1164	1	S32494	S. griseofuscus gl
	3	931.8	60.3	1167	1	N91083	DNA sequence encod
	4	903.3	58.4	2010	1	N90362	Glucose isomerase
	5	599.6	38.8	1567	1	Q01863	Sequence of the ri
	6	598.4	38.7	1185	1	Q01862	Nucleotide sequen
	7	598.4	38.7	1182	1	Q03158	Complete gene sequ
	8	481.8	31.2	1805	1	O24116	Thermus aquaticus
	9	100.8	6.5	114955	1	X53491	Human adenosine A
	10	97	6.3	1821	1	Q99366	S. lividans protea
	11	97	6.3	1821	1	N84068	Clone P5-6 encodin
	12	93	6.0	24379	1	T93095	Streptomyces freno
	13	93	6.0	24379	1	T92595	Streptomyces roseo
	14	85.2	5.5	44377	1	T78508	Streptomyces roseo
	15	85.2	5.5	44377	1	T80414	Platenolide syntha
	16	83.8	5.4	114955	1	X53491	Human adenosine A
	17	82.8	5.4	4257	1	V10362	Infected cell prot
	18	82.8	5.4	4257	1	V68520	The nucleotide seq
	19	80.6	5.2	1575	1	Q12342	Actinomycete phosph
	20	78.6	5.1	1908	1	Q99364	S. lividans protea
	21	78.6	5.1	1908	1	N84065	Tripeptidyl aminop
	22	78.6	5.1	2634	1	V22334	Microbispora therm
	23	78.4	5.1	8169	1	T26609	Actinomadura hibi
	24	77.4	5.0	30001	1	T61016	Total DNA sequenc
	25	77.4	5.0	30001	1	X05110	S. aureofaciens DNA
	26	77	5.0	833	1	Q64203	snab gene encoding
	27	77	5.0	12001	1	Q76213	HSV L/St region. H
	28	76.6	5.0	15872	1	T68715	Streptomyces venez
	29	76.4	4.9	1998	1	Q11127	Sequence encoding
	30	76.4	4.9	5392	1	Q64201	Sequence compris
	31	76.4	4.9	53789	1	V21187	Amycolatopsis med
	32	75.2	4.9	12588	1	Q63293	Sequence encoding
	33	74.6	4.8	15079	1	Q91580	S. clavuligerus cl
	34	73.6	4.8	2522	1	T87082	Hyperthermostable


```
Db 901 GCTCGGGCGGGCTGCATGCGCAACTACCTGATCTCTAAGGAGCGTGGCGGCCCTTC 960
Qy 1220 qgtgcgaccggaggttcagagagccctcgctgcgcccgcgtcgacagctcgccag 1279
Db 961 CGCGCGACCCGAGGTGAGAGAGGGGTGCGCGGTCCCGTCTGAGAGAGTGGCCCGG 1020
Qy 1280 cccacgcggcgacgagctgcagggccctgctggccgacccgacccgcgtacgagacttc 1339
Db 1021 CCACGCGGGCGAGCGTCTGCACCCCTGCTCGAGACCGGTCCGCTTCGAGAGATTC 1080
Qy 1340 gacgtgagcggcgccg--cggcgatgacctcgagcgctcgacagcctcgccatg 1396
Db 1081 GACGTGAGCGGGCGCGCGCGTGGATGGCTTCGAGCGCGCTGACAGCTGGCGATG 1140
Qy 1397 gaccactgctggcgccgcggtgga 1423
Db 1141 GACCACCTGCTGGGCGCGCGGGCTGA 1167

RESULT 4
ID N90362
AC N90362;
DF 1-NOV-1989 (first entry)
DE Glucose isomerase gene from Streptomyces spp
KW Glucose isomerase; Streptomyces; DNA.
OS Streptomyces
FH Key Location/Qualifiers
FT cds 566..1847
FT signal_peptide 566..675 /*tag= a
FT 1-NOV-1989; JP-295739.
PA (NORQ) Norinaho.
DR WPI; 89-198224/27.
DR P-PSDB; P90423.
PT Novel glucose isomerase gene
PT - exists in chromosome of Streptomyces and is
  contained in recombinant etc.
PS Claim; fig 1; pp.
CC Glucose isomerase gene from Streptomyces spp. (see P90423). Used to
  study this protein's structure by recombinant techniques, and for
  prodn. of an enzyme with improved properties.
SQ Sequence 2010 bp; 292 A; 756 C; 288 T;

Query Match 58.4%; Score 903; DB 1; Length 2010;
Best Local Similarity 81.0%; Pred. No. 4.7e-111;
Matches 1225; Conservative 0; Mismatches 265; Indels 23; Gaps 14;

Qy 2 acgagccttggtgactgggtgacaggtccacacgacacgagcgagccctcgcc 61
Db 407 ACCAGGGCCTCTGGACTGGGTGACGTGTCCACTCCGACACGA-GCGGCGCCCTCGC 465
Qy 62 tctgacatcgctctccctcttttccgctcagggctctgacctgaggttcacgc 121
Db 466 TTCTGACAT-GGGTCTCTCCCTCTCCGCGGCTCAGCCCATGTGTGGGGCGCTCC 524
Qy 122 tatgcggcgctgtggcccggggtgcggaacccgcccgttttctgttcgggt 181
Db 525 GAGGCCCCGAGGCTCCGTGGAGACATCTCTCGAAGACACCTTTGTGTGGAAGACTA 584
Qy 182 tcccttccagggagcgcgctcg---gcatactaatgttataatgccttcagcaaatagt 238
Db 585 CCTTGCTCTTCCAGAGCGGTGGTGACATCTAATTGTATAA-CGCCATGACGAATAAGT 643
Qy 239 cgaagcagcagcagcgcgcgctgactgactacacgcccaccccgagagaggttcacc 298
Db 644 CGTAGC-GAACAAGAGGAGC-GCATAGCTTCCAGCCACCCCGGAGCAAGTTACCC 701
```

```
Qy 299 ttcgccctgtgacccgtcggtgcgagggcggaaccccttcgagcagccacgcgtccc 358
Db 702 TTCGGTCTGTGAGCGTGGCTGCGAGGAAAGGACCCGTTTGGCGAGCGCCACCGCCCT 761
Qy 359 gccctgcacccggctoga---cgtgcagcggctgcccgaactggcctacggaagtacc 415
Db 762 GGCCTCGACCCGGTGCAGACCGTGCAGCGCTGGCGGAGCTGGCGCTACGGAGTGACC 821
Qy 416 ttcaacgacgacacccctgaccccttcggggcgctccgacacgagcgagggcgacgtc 475
Db 822 TTCACGACGACGACCTGAATCCCTTCGGGTGTCGCGACACCGAGCGGAGTTCGCACATC 881
Qy 476 aagcgggttcctgcagggcgtcgacgagcaccggcatgacccgttcogattggccaccaccac 535
Db 882 AAGCGTTCCGCCAGGCCCTGGACGCCACTGGCATGACGGTGGCATGGCCACCGAAC 941
Qy 536 ctcttaccaccccccgtcttcaaggcagggcggttcaacgcacacgacccgcgagtgagc 595
Db 942 CTCCTTACCCACCCCGTCTTCAAG---GACCGTTTACCCGCCAACCGCGGAGCTGGCG 998
Qy 596 cgttaacgcctgcgcaagacccatccggaacatcgatctcgcggtcgagctggggcccaag 655
Db 999 GCCTAGCGCGTGGCAAGACCATCGGCAACATCGACCTGGCGGCGGAGCTGGGCGCCAG 1058
Qy 656 gtctacgtcgctggggcgccgcgagggcgcgagtcgagtcggtccgcgaaggaagtgctg 715
Db 1059 ACGTACGTGCGCTGGGGCGCGGTGAGGGCGCGAGTCCGCTGGCGCCAGGACGAGTGCGC 1118
Qy 716 gggccctggacgcgcatagaaggagccttcgaacctgtctggcgagtaagtcacctcgag 775
Db 1119 GACGCCCTCGACCCGATGAAGGAGCGGTTGACCTCTTCGCGGAGTAGTACGTACCGCCCG 1178
Qy 776 ggctacgacatccggttcgcatcgagcccaagcgaacgagcgcgcgcgcgcgacatcgt 835
Db 1179 GGCTAGGACCTTCGCTTCGCATCGAGGCCAAGGCCAACGAGCCCCGGGGGAGATCCTC 1238
Qy 836 ctgcccacatcgccacgcctcgcttcacgagcctggagcctggagcgcgcgcgcgagctgac 895
Db 1239 CTGCCACCGTGGCCACGCCCTGGCCTTTCAGAGCGCTGGAGCGCCCGGAGCGGTGAC 1298
Qy 896 ggttcaaccccgaggtgggccaagcagagcagatggcggcctgaacctccgcacagcgcac 955
Db 1299 GCGCTCAACCCGAGGTGCGGCACGAGCAGATGGCGGCCCTGAACCTTCCCGACCGCATC 1358
Qy 956 gcgcagcctctgtggcgggcaagctcttcacatcgacctcaacgagcgcgcgcgcgcgc 1015
Db 1359 GCGCAGGCGCTGTGGCGGGCAAGCTCTTCCACATCGACCTCAACGGCGAGTCCGGCATC 1418
Qy 1016 aagtcagcaccaggaacctgcttcggcgccggtgacctgcgcgcgcgccttcgtggctgc 1075
Db 1419 AAGTACGACGAGGACTGCGG--TTCGCGCGCGGACCTGCGGGCGCGCTTCTGG-TGCTC 1475
Qy 1076 gacctgtggagagcgccggtgggaggttcgcgcacacttcgacctcaagcccccgagg 1135
Db 1476 GACCTCTCTGGAGCGCGCGGTTCAGAGGGCCCGCGGACTTTCGACTTCAAGCGCGCGGG 1535
Qy 1136 accagggacatcgacgcgctgtggcctccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1195
Db 1536 ACCGAGGACTTCGAGCGGCTGTGGCCTTCGCGGAGGCTGCATGCGCAACTACCTGATC 1595
Qy 1196 ctgaaggagcgccgcgccttcctgctgcgacccggaggttcagaggagc---cctgctg 1252
Db 1596 CTCGAAGCAGC-CGCGCGCGCTTCGCGCGCACCGGAGGTGCAGGAGCGCGCTCGGC 1654
Qy 1253 gccgcgcgcctcgacacagctcgcgagagccacccgc-ggcgcagcgctcgagggccctgct 1311
Db 1655 GCGCGCGCTCTGGACAGCTGGCCACCGGACCGGTGGCGACGCGCTGGAGGCGCTGCT 1714
Qy 1312 ggcgcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1370
Db 1715 CGCGACCGCGACCGGTTTCGAGGACTTCGAGGTGGAGCGCGCGCGCGCGCGCGATGGT 1774
Qy 1371 tcgagcgctcgacacagctcgccatggacacacccctgctggcgcccgcggtgaacccggc 1430
```


Query Match 38.7%; Score 598.4; DB 1; Length 1182;
Best Local Similarity 72.1%; Pred. No. 3e-71;
Matches 846; Conservative 0; Mismatches 301; Indels 27; Gaps 4;

QY 271 ccagccaccccccagagacagattacaccttcgacctgtgacccgtcgctgacagggcg 330
DB 9 CCAGGCGTTTCGGTGACGCCACGCAAGTTTCCTTCGGTCTCTGGAGCTTGGATGGAGGCTCG 68
QY 331 ggacccttcgagcagccacgctcccgccctcgaccggtcga---cgtgcagcgct 387
DB 69 TGACGCGTTTCGGTGACGCCACGCGTACGCACTCGACCCGCTCGAGGCGTGCACAAGCT 128
QY 388 ggcgaactcggcgccctacgagtgaccttcacacgacgacgacctgaccccttcggg 447
DB 129 CGCTGAGATCGCGCGCTACGGCATCAGTTCCACGACGACGACCTGCTGCTTCGGCTC 188
QY 448 gtccacacccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 507
DB 189 GGAGCGCCACAGACCGCGGATCGGATCATCGGGCTTCAAGAGGGCGCTCGAGGACCGG 248
QY 508 catgacgcttcgagtcgac 567
DB 249 CCGTATCGTCCGATGGTGACCCACCACTCTTACCCACCCCGGTGTTCAAGGACGCGG 308
QY 568 gttacccgcaacgacccgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 627
DB 309 CTTACACGACACGACCGTTCGGTGGCGGTACGGCATCCGCAAGGTGCTGCGCCAGAT 368
QY 628 cgaatcgcggtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 687
DB 369 GGACCTTCGGCGCGAGTGGGCGGAGAGGCTGCTCTCTGGGGCGCGCGAGGGCGC 428
QY 688 ggagtcggtgcgcaagagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgag 747
DB 429 CGAGTACGACTCGGCGAGGCTCAGCGCGCGCTCGACCGCTACCGCGAGGCGCTCAA 488
QY 748 cctgctcggcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgag 807
DB 489 CCGTCTCGCGCGAGTACTCCGAGGACCGCGGTACGGCTTCGGCTTCGCCATCGAGCGAA 548
QY 808 ggcgaacgagcgcgagcgcgagcgcgagcgcgagcgcgagcgcgagcgcgagcgcgag 867
DB 549 GCGAAGAGCGCCCGGCGAGATCTGCTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 608
QY 868 cgagtcgctggagcgcgagcgcgagcgcgagcgcgagcgcgagcgcgagcgcgagcgc 927
DB 609 GCAGGAGCTGGAGCGTCCGAGCTCTCGCATCAACCCGAGACCGGCGACGAGCAGAT 668
QY 928 ggcggcctgaactccgacgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 987
DB 669 GTCGAAGCTCAACTTCAACCGAGGATCGCCAGGCGCTGTGCAAGAAGCTGTTCGA 728
QY 988 catgacacacgagcgcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 1047
DB 729 CATCGACCTGAAGCTGACGACGCGCGGAGGCTTCGACGAGGAGCTGCTTCGGCCACG 788
QY 1048 tgacctgcgagcgcgagcgcgagcgcgagcgcgagcgcgagcgcgagcgcgagcgcgag 1092
DB 789 TGACCTGCTCAACGCGTCTCGCTGTGCGACCTCTTGAGACGCGCGGCGGCGCGCC 848
QY 1093 cggctcggggggttcgagcagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 1152
DB 849 GCGGTACGACGAGCGCGCTCACTTCGACTACAGCGCGTCCGCTACCGAGGACTACGACG 908
QY 1153 cgtgtgggctcggcgagcgcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 1212
DB 909 CCGTCTGGGAGTCGGCGAAGGCCAACATCCGATGTGTACCTGCTCTCAAGAGCGGCGCAA 968
QY 1213 cgccttcgagtcgagcgcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 1272
DB 969 GCGGTTCGCGCGACCGCGAGGTGAGGAGGCGCTTCGCGCGCGACGAGGTCGCGAGCT 1028

QY 1273 -----cgccgagccacccgagcgcgagcgcgagcgcgagcgcgagcgcgagcgcg 1326
DB 1029 GAAGACCCCGACCCCTGAACCCCGGCGGAGGATACCCGAGCTGCTCGCGACCGACGC 1088
QY 1327 gtacgagacactcagctgagcgc---ggccgcgcgcgagcgcgagcgcgagcgcgagcgc 1383
DB 1089 GTTCGAGGACTACGACGCGGACCCGCTGGGCGCCAGGGCTTCGGCTTCGCTCAAGCTGA 1148
QY 1384 ccagctcgcacatggaccacacacacacacacacacacacacacacacacacacac 1417
DB 1149 CCAGCTCGGATCGAGCAGCTGCTCGGAGCGCGC 1182

RESULT 8
Q24116
ID Q24116 standard; DNA; 1805 BP.
AC Q24116;
DT 03-NOV-1992 (first entry)
DE Thermus aquaticus xylose isomerase gene.
KW Large scale; recombinant; colourless; fructose; ss.
OS Thermus aquaticus.
FH Key Location/Qualifiers
FT cds 288..1461
FT /*tag= a
FN EP-483691-A.
PD 06-MAY-1992.
PF 25-OCT-1991; 118264.
PR 07-MAR-1991; JP-291067.
PR 07-MAR-1991; JP-067967.
PR 24-JUN-1991; JP-178598.
PA (UDAK)/ UDAKA S. Sakaguchi K, Uda S, Yamagata H;
PI Dekker K, Sakaguchi K, Uda S, Yamagata H;
DR WPI: 92-152378/19.
DR P-PSDB; R2623.
PT Xylose isomerase gene of Thermus aquaticus and its encoded
PT protein - expressed on large scale and used to produce conc.
PT fructose from glucose
PS Claim 1; Fig 1-3; 24pp; English.
CC To obtain the DNA sequence of the xylose isomerase gene from Thermus
CC aquaticus genomic DNA from T. aquaticus was digested by SacI and
CC fractionated by agarose gel electrophoresis. The DNA was denatured
CC and the gel neutralised. The DNA fragments were blotted onto a
CC nylon membrane and hybridised with a labelled 287 base Bgl fragment
CC of a Streptomyces griseofuscus S-41 xylose isomerase gene. A single
CC band of 1.8 kb was extracted with glassmilk and ligated into pUC18
CC digested with SacI. This was used to transform E. coli JM109 and
CC transformants selected using the above probe. A 1.8 kb SacI
CC fragment which hybridised strongly to the probe was cleaved into
CC smaller fragments and subcloned into E. coli JM109 using pUC18 and
CC pUC19. These fragments were then sequenced. The gene allows large
CC scale expression of T. aquaticus xylose isomerase which is used in the
CC prodn. of fructose at higher temperatures than those previously
CC possible. Xylose isomerase is used in the prodn. of relatively
CC colourless and highly conc. fructose which is free of coloured
CC byproducts, such as psicose, usually formed when the pH of the reaction
CC mixture exceeds pH 7.
SQ Sequence 1805 BP; 295 A; 558 C; 627 G; 325 T;

Query Match 31.2%; Score 481.8; DB 1; Length 1805;
Best Local Similarity 62.5%; Pred. No. 4.6e-56;
Matches 788; Conservative 0; Mismatches 467; Indels 6; Gaps 2;

QY 269 taccagccaccccgagcagagtcacacacacacacacacacacacacacacacacacac 328
DB 291 TAGAGCCCAACCGGAGCAGAGTTTACCTTGGCTTGGACTTGGGCAATGGGCG 350
QY 329 cgggaccccttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 385
DB 351 CGTGATCCCTTCGGGAGCGCGTTCGGGAGAGGCTGGAGCCCGGTTTACGTGTTTCAAG 410
QY 386 ctggcgcgaactggcgccctacagtgacctccacgacgacgacgacgacgacgacgacg 445

QY 214 ttgtaaatccctcgtacgacaaatagtcgcaagcgagcaagagccggtgcatgaactacca 273
DB 104902 VGGCCVGGCGNNHNNNSGCGCGNNHNNNSVGGCGNNHNNNSVGGCGVGGCG 104843
QY 274 gccacccccaggagacagttcaccttcgctctgtagacgtcgagcgagggcgga 333
DB 104842 NHHNNNSCCVGGCCVGGCGNNHNNNSCCVGGCCVGGCGNNHNNNSGCGCCVGGCG 104783
QY 334 ccccttcggcgagcgacgcacgtcccgccctcagccggtcgagctgagcgagcggtgcccga 393
DB 104782 NHHNNNSCCVGGCCVGGCGNNHNNNSGCGCCVGGCGVGGCGNNHNNNSGCGCCV 104723
QY 394 actggtgctacgagtgatgacattccacagacagacacgtatcccccttcggggcgctcca 453
DB 104722 GGCVCVGGGNNHNNNSGCGCGCCVGGCGVGGCGNNHNNNSGCGCGCCVGGCGV 104663
QY 454 caccgagcgagggcgacgtcaagcggttcctcagcgctcgagcgctcgagcgacgcgcatgac 513
DB 104662 NHHNNNSCGCGCGCCVGGCGVGGCGNNHNNNSGCGCGCGCCVGGCGVGGCGNNHNN 104603
QY 514 cgttcgatggccac 573
DB 104602 NSGGCGCGCGCCVGGCGVGGCGNNHNNNSGCGCGCGCCVGGCGVGGCGNNHNN 104543
QY 574 cgcacacagcgagcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 633
DB 104542 SCCGCGCGCGCGCCVGGCGVGGCGNNHNNNSGCGCGCGCCVGGCGVGGCGNNH 104483
QY 634 cgcgttcgagtcggcgacagtcacgtcgtcgttcgttcgttcgttcgttcgttcgttcgttc 693
DB 104482 NHHNSGCGCGCGCGCCVGGCGVGGCGNNHNNNSGCGCGCGCCVGGCGVGGCGNNH 104423
QY 694 cgttcgacgagcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 753
DB 104422 GCGGNNHNNNSCGCGCGCGCGCCVGGCGVGGCGNNHNNNSCGCGCGCGCGCGCGCG 104363
QY 754 cgcgagtcagtcacgtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 813
DB 104362 CGCCVGGCGVGGCGNNHNNNSGCC--CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104305
QY 814 cgcgagcgcgcgagac 873
DB 104304 SGCG 104245
QY 873 ---gcttcgagcgccccagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 929
DB 104244 CVGGCGVGGCGNNHNNNSGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104185
QY 930 cggcgtgacattccgcagcgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 989
DB 104184 CGGACNNHNNNSGCG 104125
QY 990 tcgacacacagcgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 1049
DB 104124 CGGACCGNNHNNNSCG 104065
QY 1050 acttcgagcgcccccttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 1109
DB 104064 NSGCCCG 104005
QY 1110 gcaacttcagtcac 1169
DB 104004 CGGACCG 103945
QY 1170 cgggttcgacgcaactac 1229
DB 103944 GCGCGCAVGGNNHNNNSGCG 103885
QY 1230 cggaggttcagaggttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 1289
DB 103884 CGGCCAVGAGNNHNNNSGCG 103825

QY 1290 ccgacggcctcagagccctcgtggtgacgacgacgacgacgacgacgacgacgacgacgacg 1349
DB 103824 AGCGCGCAVAGACGCGNNHNNNSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 103765
QY 1350 cggc 1409
DB 103764 CGGCCCGAGCGCGCAVAGACGCGCGNNHNNNSGCGCGCGCGCGCGCGCGCGCGCGCG 103705
QY 1410 gcycccgcggtgaaaccg 1427
DB 103704 CNHNNNSGCGCGCGCG 103687
RESULT 10
QY9366
ID Q99366 standard; DNA; 1821 BP.
AC Q99366;
DC 04-DEC-1995 (first entry)
DE S. lividans protease P5-6 gene.
KW Protease; metalloprotease; tripeptidyl aminopeptidase;
KW Protease-deficiency; protein secretion; ds.
OS Streptomyces lividans.
FH Key Location/Qualifiers
FT cds 104..1723
FT signal_peptide 104..244
FT /tag= a
FT /tag= b
FT mat_peptide 245..1720
FT /tag= c
FT /product= P5-6.
FT W09517512-A.
PN 29-JUN-1995.
PF 22-DEC-1994; U14772.
PR 23-DEC-1993; US-173508.
PA (CANG-) CANGENE CORP.
PI Bartfield D, Butler MJ, Hadary D, Jenish DL, Krieger TJ;
PI Malek LT, Soostmeyer G, Walczyk E;
DR WPI: 95-240673/31.
DR P-PSDB; R80508.
PT Endogenous Streptomyces protease(s), opt. having impaired activity -
PT useful in prodn. of exogenous proteins with reduced proteolytic
PT degradation
PS Claim 7: Fig.14: 142pp; English.
CC Protease Tap-negative cells were transformed with a S. lividans 66
CC genomic library and screened with APA-beta-naphthylamide to
CC isolate colonies contg. genes (Q99365-68) for novel proteases
CC P5-4, P5-6, P5-10 and P8-2 (R80505-08). Impaired expression of
CC such proteases by Streptomyces hosts improves the quality, quantity
CC and stability of exogenous gene products.
SQ Sequence 1821 BP; 283 A; 715 C; 602 G; 221 T;

Query Match 6.3%; Score 97; DB 1; Length 1821;
Best Local Similarity 44.8%; Pred. No. 7.7e-06;
Matches 529; Conservative 0; Mismatches 625; Indels 27; Gaps 3;
QY 240 gcaagcgagcaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 299
DB 543 GCTACCGCGGAAGTGTGCG 602
QY 300 tcggcgtgtgacgtcggttcgagggcgagggcgagggcgagggcgagggcgagggcgagggcg 359
DB 603 CCCGAGTGAACCCCTCGAGTGCCTGGACGGCGCGGAGATGGACGGTACACGCGGCGCGCG 662
QY 360 ccttcgacccggttcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgac 419
DB 663 AGTCACCCCGACGACG 722
QY 420 acgagcgagcgtatcccccttcggggcggtccgacacgagcgagcgagcgagcgagcgagcgagc 479
DB 723 CCGAGGG-----CTGCGGGGCGGACGCGCGGAGGTGTGCGCCACGCTCCA 770
QY 480 gtttcgtcagcgctcgc 539

(HOFF) HOFFMANN LA ROCHE & CO AG F.
Reeves CD, Soliday CL;
WPI; 97-538619/50.
P-PSDB; W34199-W34219.
Streptomyces frenolicin gene cluster - useful for producing recombinant frenolicin antibiotics

Claim 1: Page 40-60; 6pp; English.
This DNA sequence comprises the Streptomyces frenolicin gene cluster containing specifically claimed coding sequences (genes A-U) that respectively encode 21 proteins (see W34199-219) involved in frenolicin synthesis. The genes can be divided into 5 subclusters: (1) genes A, B, C, D and F encode an efflux pump; (2) genes H, I, J and K encode butyrate starter synthases; (3) genes L, M and N encode polyketide synthases (PKS); (4) genes O, P, Q and R encode a hemiketalase, a ketoreductase and cyclases/dehydrases; and (5) genes S and T encode a keto/enoyl reductase and a hydrolase. Also claimed are vectors, host cells (especially a Streptomyces sp., particularly Streptomyces roseofulvus), and the encoded proteins. Cells transformed using the above sequence can be cultured to produce frenolicins or frenolicin precursors. The precursors can be converted to frenolicins by chemical or other methods. The frenolicins can be oxidised to frenolicin B, an antibiotic used as an anticoccidial agent. The frenolicins can be used as animal feed additives.

Sequence 24379 BP; 3077 A; 9792 C; 8499 G; 3011 T;

Query Match 6.0%; Score 93; DB 1; Length 24379;
Best Local Similarity 44.5%; Pred. No. 1.5e-05;
Matches 415; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

QY 484 ccgtcaggcgctcgaagcgacgcgcatcacgcttccatggccaccaccaactcttcac 543
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15033 CCGGCTCGGCTCTCGCCGCCAGGACACTGAAGACGCCGCCTGACC CGCGCGC 15092
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 544 ccaccccgtttcaagcgagcgcggttcaccgcccaacgccgcgcagtgcgcgttacgc 603
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15093 GCTCGCCGAGTACGGGGCGAGCGCGCTCACCTCAACGCCACCGCGGTTCAGTTTAC 15152
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 604 cctgcgaagaacctcoggaaac---atgatctcgcggtcgagctggcgccaagtcta 660
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15153 CCACCGGAGATCCGCAAACTTGACCGAGGGCCCCCGCGCTACGCGTCTACGAGTC 15212
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 661 cgctgcctggggcgcgcgagggcgcgagtcoggtgcgscgaagacgtgctgcgcgc 720
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15213 CTTCGCTGGTCTACGCCGTCAACACGCCAGATCTCCATCCGGCAGGATCGCGG 15272
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 721 cctggacgcgatgaagagggccttgacacctgctgcgagtagctgcacctgcagggcta 780
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15273 CCCCGGCGCGCTCGTGCGCGACACGAGCGGGCGCGCTCGACGCCCTCGCCAGGCGCG 15332
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 781 gcacatcgggttcgcatcgacccaacgaacgcgcgcgcgcgcacatcgtcgc 840
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15333 CCGCGTACTGCGCAAGGCGGGGTGCTGGCGGTGAGCGGCGGTGGAGTCCGCGCTCGA 15392
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 841 caccatcgcccaacgcgtcgccttcacatcgagcgcttgagcgccccgagctgtaagggt 900
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15393 CCCCCTGGGCGCTGGCGGCCACAGCGCTCGTCGGGCACCTCAGCCGCTCCGGCGACCGCGC 15452
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 901 caaccccgagtgggccacgcagatggccggcctgaactcccgacggcatcgaca 960
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15453 CACCGCTACCTCCCTTCGACCGCGGTGCTCGCGCACCGTGTGCGGAGGCGGCGC 15512
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 961 ggctctgggggggaagctcttcacatcgacctaacgcgcagtcggcataagta 1020
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15513 CCTCCTCACCTGAGACACCCCGGACCGCGAGGAGGCGGACACCCCGATTACGG 15572
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1021 cgacccagacctgcgttcgcgccggtgacctgcgcgcgcctcttgctgtgacct 1080
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 15573 CGAACTCCCGGGGTACGCCGGCCACTTCGACCCGCGCGGGGCTCGGNAGGCCGCCCGC 15632
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1081 gctggagagcgcgctgggggggtccgcgcacattcgacttaagcccccgacgga 1140
| | | | | | | | | | | | | | | | | | | | | | | | | | |

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2000, 23:05:43 ; Search time 44.32 Seconds
(without alignments)
4176.263 Million cell updates/sec

Title: US-09-383-318-1
Perfect score: 1546
Sequence: 1 caccagcgccttggtgact.....gcgcggtgtggggcgctgc 1546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 214294 seqs, 59861574 residues

Total number of hits satisfying chosen parameters: 428588

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PTUS9_COMB.seq.*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	952.8	61.6	1164	1 US-07-640-476-6	Sequence 6, Appli
2	599.6	38.8	1566	7 5290690-3	Patent No. 5290690
3	598.8	38.7	1185	7 5290690-1	Patent No. 5290690
4	598.4	38.7	1182	1 US-07-640-476-4	Sequence 4, Appli
5	481.8	31.2	1805	1 US-08-112-630-1	Sequence 1, Appli
6	98.6	6.4	1820	1 US-08-173-508-7	Sequence 7, Appli
7	97	6.3	1821	3 US-08-265-310-7	Sequence 7, Appli
8	85.2	5.5	44377	3 US-08-804-227C-7	Sequence 7, Appli
9	85.2	5.5	44377	4 US-08-804-198-1	Sequence 1, Appli
10	82.8	5.4	4257	3 US-08-890-473-1	Sequence 1, Appli
11	79	5.1	1668	4 US-08-901-547A-1	Sequence 1, Appli
12	78.6	5.1	1908	1 US-08-173-508-1	Sequence 1, Appli
13	78.6	5.1	1908	3 US-08-265-310-1	Sequence 1, Appli
14	78.6	5.1	12001	2 US-08-458-568A-11	Sequence 11, Appli
15	77.4	5.0	30001	1 US-08-125-468-1	Sequence 1, Appli
16	77.4	5.0	30001	3 US-08-474-933-1	Sequence 1, Appli
17	77	5.0	1620	3 US-08-461-775-10	Sequence 10, Appli
18	76.4	4.9	1998	7 5212296-8	Patent No. 5212296
19	76.4	4.9	5392	3 US-08-403-852D-1	Sequence 1, Appli
20	75.4	4.9	833	3 US-08-403-852D-3	Sequence 3, Appli
21	75.2	4.9	12588	4 US-08-387-942C-1	Sequence 1, Appli
22	73.6	4.8	2522	1 US-08-559-260-1	Sequence 1, Appli
23	72.2	4.7	2668	3 US-08-461-775-11	Sequence 11, Appli
24	72	4.7	1931	4 US-09-130-114-2	Sequence 2, Appli
25	72	4.7	20235	2 US-07-642-734C-3	Sequence 3, Appli
26	71.2	4.6	8438	1 US-07-945-283-1	Sequence 1, Appli
27	71	4.6	1146	2 US-08-482-385A-1	Sequence 1, Appli

28	71	4.6	2728	2 US-08-482-385A-5	Sequence 5, Appli
29	70.4	4.6	1268	3 US-08-403-852D-2	Sequence 2, Appli
30	70.4	4.6	2312	1 US-07-736-178C-1	Sequence 1, Appli
31	70.4	4.6	43280	3 US-08-804-227C-1	Sequence 1, Appli
32	70	4.5	2721	7 5215881-2	Patent No. 5215881
33	69.2	4.5	1208	3 US-08-403-852D-4	Sequence 4, Appli
34	68.8	4.5	1771	3 US-08-533-669A-7	Sequence 7, Appli
35	68.8	4.5	1771	4 US-08-511-872-1	Sequence 1, Appli
36	68.8	4.5	11219	2 US-07-642-734C-1	Sequence 1, Appli
37	68.6	4.4	3468	1 US-07-951-715A-2	Sequence 2, Appli
38	68.6	4.4	3468	3 US-08-459-449A-2	Sequence 2, Appli
39	67.2	4.3	4257	3 US-08-690-473-1	Sequence 1, Appli
40	66.2	4.3	1288	2 US-08-440-856A-9	Sequence 9, Appli
41	66	4.3	2064	1 US-08-343-428-1	Sequence 1, Appli
42	65.6	4.2	30001	1 US-08-125-468-1	Sequence 1, Appli
43	65.6	4.2	30001	3 US-08-474-933-1	Sequence 1, Appli
44	64.4	4.2	4776	4 US-08-852-401-1	Sequence 1, Appli
45	63.8	4.1	8051	4 US-08-576-626A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-07-640-476-6
; Sequence 6, Application US/07640476
; Patent No. 5376536
; GENERAL INFORMATION:
; APPLICANT: QUAX, WILHELMUS
; APPLICANT: LUITEN, RUDOLF G.M.
; APPLICANT: SCHUURHUIZEN, PAUL W.
; APPLICANT: MRABET, NADIR
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND
; TITLE OF INVENTION: THEIR USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,476
; FILING DATE: 19910110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kate H. Murashige
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20009.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces murinus
; STRAIN: DSM 40091
; FEATURE:

	1089	1384	1149
Db	1089 GTTTCAGGACTACGACGCGACCGCTGGCGCCAAAGGCTTCGCTCAAGCTGAA	1384 CGAGCTCGCATGGACCACTGCTGGGCGCCCGC	1149
Qy			
Db			

RESULT 5

```

US-08-112-630-1
: Sequence 1, Application US/08112630
: Patent No. 5411886
: GENERAL INFORMATION:
: APPLICANT: UDAKA, Shigezo
: APPLICANT: SAKAGUCHI, Kenji
: APPLICANT: YAMAGATA, Hideo
: APPLICANT: DEKTER, Koen
: TITLE OF INVENTION: XYLOSEISOMERASE GENE OF THERMUS
: TITLE OF INVENTION: AQUATICUS, XYLOSEISOMERASE AND PROCESS FOR PREPARATION OF
: TITLE OF INVENTION: FRUCTOSE
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Burns Doane, Swecker and Mathis
: STREET: George Mason Bldg., Washington & Prince Sts.
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/112,630
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/783,150
: FILING DATE: 28-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Rea, Teresa S.
: REGISTRATION NUMBER: 30,427
: REFERENCE/DOCKET NUMBER: 024705-007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: TELEX: 440 580
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1805 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 288..1448
: US-08-112-630-1

```

[illegible]

Db 1491 AGGCCCTGGTCTGGACGAGGAGGGTAGAAAGCGCGTGAAGCGCGGCGCGGTACCCC 1550
Qy 1523 c 1523
Db 1551 c 1551

RESULT 6

US-08-173-508-7
; Sequence 7, Application US/08173508
; Patent No. 5618485
; GENERAL INFORMATION:
; APPLICANT: Bartfield, Daniel
; APPLICANT: Butler, Michael J.
; APPLICANT: Hadary, Dany
; APPLICANT: Jenish, David
; APPLICANT: Krieger, Timothy
; TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
; TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,508
; FILING DATE: 23-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 18740/125/CACO
; TELEPHONE: 202 672 5300
; TELEFAX: 202 672 5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1820 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 104..1720
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 104..244
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 245..1720
; US-08-173-508-7

Query Match 6.4%; Score 98.6; DB 1; Length 1820;
Best Local Similarity 44.2%; Pred. No. 3.6e-09;
Matches 572; Conservative 0; Mismatches 694; Indels 27; Gaps 3;
Qy 240 gcaagcgagcaagagcgccggcatgaactaccagccaccaccccgagagacaggttcaact 299
Db 543 GCTACCCGGCGAAGTCCGCGCCACAGTACACATGGTGGCGGTGACACCCCGGGGCGTGG 602
Qy 300 tcggcctgtgagaccctgcgtgagggcgggagcccttcggcgagccacccggtcccg 359

Db 603 CCCACGTGAACCCGTCGAGTGCCTGGACGGGCGCGAGATGGACGCGTACACGCGCACCG 662
Qy 360 cctcgcaccggtcgagctgcagcggtgcccgaactggcgccctacgagtgaccttcc 419
Db 563 AGTCAACCCCGGACGCGGGCGAGACGAGAGCTGTGTGAGCGCTCAAGAGAGTTCG 722
Qy 420 acgacgacgacgtatcccttcggcggtccgacacacgagcgagcgacgctcaagc 479
Db 723 CCGAGGG-----CTGGGGGCGGAGCGCGCGAAAGCTGCTGCGCCACGCTCCTCA 770
Qy 480 gtttcgtagggcgctcgacgacccggcatgacggttccgagtggccaccacacacctct 539
Db 771 CGTTCGAGGGCGGACGCGGACATGACGTCTGTGCGCGGTGCTGGGCGACGAGAAGTGA 830
Qy 540 tcaccaccccgctcttcaaggcggttcacgcggttcacgcgcaaacgacgcgcgagtcgcggt 599
Db 831 CCTACGTGGGAGCGTCTGACGGCACCTTCTTGGCGGGGACCTACCGCGGTCTGTTCCTCCG 890
Qy 600 acgcccctgcgaagaccatccgggaacatctctcgcggtcgagctggcgcccaaggtct 659
Db 891 ACCGACGGGCGCGCTGGTCTCTGGAC-----GGCGCATGACACCCCTCGCTGC 938
Qy 560 acgtcgctggggcgcgcgagggcgcgagtcgggtgcgcgcaagagacgtgcgtgcgg 719
Db 939 CCGCCCGCCCGCTGAACCTGGAGCAGACGAGGGCTTCGAGACGGCGTTCCAGTCTCTCG 998
Qy 720 ccttggaccgcataagagggccttcgacctgctcgcgagtagctcacctcgagggct 779
Db 999 CGAAGGACTGGTGAAGCAGCCGACTGCCCTTCGGCGACAGGACACACCC---CCG 1055
Qy 780 acgacatcgtgtgcacatcgagcccaagcgaacgagcgcgcgcgacatcctgtgc 839
Db 1056 ACCAGGTGCGGAAGAACTCAAGTCTTCTTGGACGACCTTGGACGCGAAGCGCCCTGCGCG 1115
Qy 840 ccaccatcgccacgcgctcgcttcgagcgcttcgagcgcccgagcgagctgacggtg 899
Db 1116 CCGCGACGCGCGCGCGCGCAAGCTACCCGAAATCCCTTCGCCACACCGCGGTGTATCGCG 1175
Qy 900 tcaaccccgaggtggcgacgagcagatggcgcgctgaacttcgacgacgagtcgcgc 959
Db 1176 CGATGTACGACGAGGGCGCTGGGACGAGCGAGTGGCGGAGTCCCTACCTCGCGGATCAAG 1235
Qy 960 aggtctgtggcggaagctcttcacatcgacctcaacggcgagtcggcgacatcaagt 1019
Db 1236 AGAAGGACGTTGCGGGCTGCTGATCTCTCCGACAGCTACTACGAGCGGCGAGCGCAGC 1295
Qy 1020 acgacagggacctgcgttcggcgcggtgacctgcgcgcgcttcgtggtgcacc 1079
Db 1296 GCGGCTACAGCAACCTGATGTTCCCAACGCGCGCTGAACCTGCTCGACCTCCCGCGCG 1355
Qy 1080 tgcctggagagcgcggtgggaggggtccgcgcacttcgacctcaagcccccgcgagcg 1139
Db 1356 CTTCTCTCTCCCGACGAGGTGGCGACGCGCCCTCCCGACTTCGAGAGGCGTCCCGCG 1415
Qy 1140 aggaatcgagcggtgtgggcttcggcgcggtgcatgcgaactacatgacctga 1199
Db 1416 TCTTCGGGAGGGCTCGCCTGGTCTCTCCCTGAACTACGCGTACTGCGCGTGAAGCCCA 1475
Qy 1200 aggagcgcgcgcgcttcgctcgacaccccgaggttcagggagggccttcggtgcgcc 1259
Db 1476 CGGGGAGCGCGCACCGCATCGAGCGCGCGCGCCACCGCGCATCGTGGTTCGGCACCA 1535
Qy 1260 ggcctcgacagctcgcgagaccccgcgcgcgagtcgagcggttcgctgctgctgcgac 1319
Db 1536 CCGCGGACCGGCGCACCCCTACCGCTGGCGCGAGGGCCCTTCGACACGAGTCACTCCG 1595
Qy 1320 gcacgcgtacgaggaacttcgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1379
Db 1596 GCCACCTCTCACCTACGAGGAGAGCGGCACACCGCTACGCGCGCGGCGGACGCTCTGCA 1655
Qy 1380 tcgacagctcgccatgagacacacctgctggcgcccggtgacacggcgagcgagggg 1439
Db 1656 TCGACTCCGGGATCAACACGTACTCTGTCGACCGCGACCGCGCGCGGACGCGGACGCGCT 1715

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2000, 23:04:23 ; Search time 513.23 Seconds
(without alignments)
11373.412 Million cell updates/sec

Title: US-09-383-318-1
Perfect score: 1546
Sequence: 1 cacgagccttggtgact.....gcgcggtgtggggcgctgc 1546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4538634 seqs, 1887831982 residues

Total number of hits satisfying chosen parameters: 9077268

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST.*
1: em_est1.*
2: em_est2.*
3: em_est3.*
4: em_est4.*
5: em_est5.*
6: em_est6.*
7: em_est7.*
8: em_est8.*
9: em_est9.*
10: em_est10.*
11: em_est11.*
12: em_est12.*
13: em_est13.*
14: em_est14.*
15: em_est15.*
16: em_est16.*
17: em_est17.*
18: em_est18.*
19: em_est19.*
20: gb_est1.*
21: gb_est2.*
22: gb_est3.*
23: gb_est4.*
24: gb_est5.*
25: gb_est6.*
26: gb_est7.*
27: gb_est8.*
28: gb_est9.*
29: gb_est10.*
30: gb_est11.*
31: gb_est12.*
32: gb_est13.*
33: gb_est14.*
34: gb_est15.*
35: gb_est16.*
36: gb_est17.*
37: gb_est18.*
38: gb_est19.*
39: gb_est20.*
40: gb_est21.*
41: gb_est22.*
42: gb_est23.*
43: gb_est24.*
44: gb_est25.*

45: gb_est26.*
46: gb_est27.*
47: gb_est28.*
48: gb_est29.*
49: gb_est30.*
50: gb_est31.*
51: gb_est32.*
52: em_est20.*
53: em_est21.*
54: em_est22.*
55: em_est23.*
56: em_est24.*
57: em_est25.*
58: em_est26.*
59: gb_est33.*
60: gb_est34.*
61: gb_est35.*
62: gb_est36.*
63: gb_est37.*
64: gb_est38.*
65: em_est27.*
66: em_est28.*
67: em_est29.*
68: em_est30.*
69: gb_est39.*
70: gb_est40.*
71: gb_est41.*
72: gb_est42.*
73: gb_est43.*
74: gb_est44.*
75: em_est31.*
76: em_est32.*
77: em_est33.*
78: em_est34.*
79: gb_gss1.*
80: gb_gss2.*
81: gb_gss3.*
82: gb_gss4.*
83: em_gss1.*
84: em_gss2.*
85: em_gss3.*
86: em_gss4.*
87: gb_gss5.*
88: gb_gss6.*
89: gb_gss7.*
90: gb_gss8.*
91: gb_gss9.*
92: em_gss5.*
93: em_gss6.*
94: em_gss7.*
95: em_gss8.*
96: em_gss9.*
97: em_gss10.*
98: em_gss11.*
99: gb_gss10.*
100: gb_gss11.*
101: em_gss12.*
102: gb_gss12.*
103: gb_gss13.*
104: gb_gss14.*
105: gb_gss15.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	82	5.3	925	79	CNS0091P	AL053013 Drosophil
2	78.8	5.1	935	79	CNS006XK	AL066031 Drosophil

Db 394 GTGKNTTTCVGTCTTKSSCTBSSGTSBSTS 361

Search completed: March 22, 2000, 23:16:27
Job time: 724 sec

THIS PAGE BLANK (85-70)

[illegible]

```

RESULT      2
US-07-637-399-7
; Sequence 7, Application US/07637399
; Patent No. 5340738
; GENERAL INFORMATION:
; APPLICANT: Lambeir, Anne-Marie
; APPLICANT: Lasteris, Ignace
; APPLICANT: Mrabet, Nadir
; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Van Der Lean, Jan M.
; APPLICANT: Misset, Onno
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN
; TITLE OF INVENTION: ALTERED PH OPTIMUM
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/637,399
; FILING DATE: 19910104
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 4615-0020.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-07-637-399-7

```

```

1  RESULT 3
2  US-07-640-476-7
3  ; Sequence 7, Application US/07640476
4  ; Patent No. 5376536
5  ; GENERAL INFORMATION:
6  ; APPLICANT: QUAX, WILHELMUS
7  ; APPLICANT: LUITEN, RUDOLF G.M.
8  ; APPLICANT: SCHUURHUIZEN, PAUL W.
9  ; APPLICANT: MRABET, NADIR
10 ; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND
11 ; TITLE OF INVENTION: THEIR USE
12 ; NUMBER OF SEQUENCES: 22
13 ; CORRESPONDENCE ADDRESS:
14 ; ADDRESSEE: Morrison & Foerster
15 ; STREET: 545 Middlefield Road, Suite 200
16 ; CITY: Menlo Park
17 ; STATE: CA
18 ; COUNTRY: USA
19 ; ZIP: 94025
20 ; COMPUTER READABLE FORM:
21 ; MEDIUM TYPE: Floppy disk
22 ; COMPUTER: IBM PC compatible
23 ; OPERATING SYSTEM: PC-DOS/MS-DOS
24 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
25 ; SOFTWARE:
26 ; CURRENT APPLICATION DATA:
27 ; APPLICATION NUMBER: US/07/640,476
28 ; FILING DATE: 19910110
29 ; CLASSIFICATION: 435
30 ; ATTORNEY/AGENT INFORMATION:
31 ; NAME: Kate H. Murashige
32 ; REGISTRATION NUMBER: 29,959
33 ; REFERENCE/DOCKET NUMBER: 24615-20009.20
34 ; TELECOMMUNICATION INFORMATION:
35 ; TELEPHONE: (415) 327-7250
36 ; TELEFAX: (415) 327-2951
37 ; TELEX: 706141
38 ; INFORMATION FOR SEQ ID NO: 7:
39 ; SEQUENCE CHARACTERISTICS:
40 ; LENGTH: 388 amino acids

```

TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-640-476-7

Query Match 93.6%; Score 1900; DB 1; Length 388;
Best Local Similarity 92.8%; Pred. No. 2.7e-182;
Matches 360; Conservative 14; Mismatches 12; Indels 2; Gaps 2;

QY 1 MNYQPTPEDRFTGLTWGQGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
DB 1 MSQPTPEDRFTGLTWGQGRDPFGDTRPALDPVETVORLAELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVRRFRQALDQATGTVPMATTNLTTHPVFKAGAFANDRAVRYALRKTII 119
DB 61 FGSSDTERESHKFRFRQALDQATGTVPMATTNLTTHPVFKDGGFTANDRDRVRYALRKTII 120
QY 120 RNIDLAELGAKYVAVWGREGAESGAADKDVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
DB 121 GNIDLAELGAKYVAVWGREGAESGAADKDVRAALDRMKEAFDILLGEYVTAQGYDLRFAI 180
QY 180 EKPNEPRGDILLPTTGHALAFIERLERPELYGVNPEVGEQWAGLNFPHGIAQALWAGK 239
DB 181 EKPNEPRGDILLPTTGHALAFIERLERPELYGVNPEVGEQWAGLNFPHGIAQALWAGK 240
QY 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGVW 299
DB 241 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGVW 300
QY 300 ASAAGCMRNYLILKERAAPRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDF 359
DB 301 ASAAGCMRNYLILKERAAPRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDF 360
QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
DB 361 DVDAARGMAFERLDQALMDHLLGARG 388

RESULT 4
US-08-112-703-7
Sequence 7, Application US/08112703
Patent No. 5384257

GENERAL INFORMATION:
APPLICANT: Lambel, Anne-Marie
APPLICANT: Lasters, Ignace
APPLICANT: Mrabet, Nadir
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Van Der Laan, Jan M.
APPLICANT: Misset, Onno
TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN
TITLE OF INVENTION: ALTERED PH OPTIMUM
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/112,703
FILING DATE: 26-AUG-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 4615002001

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-112-703-7

Query Match 93.6%; Score 1900; DB 1; Length 388;
Best Local Similarity 92.8%; Pred. No. 2.7e-182;
Matches 360; Conservative 14; Mismatches 12; Indels 2; Gaps 2;

QY 1 MNYQPTPEDRFTGLTWGQGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
DB 1 MSQPTPEDRFTGLTWGQGRDPFGDTRPALDPVETVORLAELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVRRFRQALDQATGTVPMATTNLTTHPVFKAGAFANDRAVRYALRKTII 119
DB 61 FGSSDTERESHKFRFRQALDQATGTVPMATTNLTTHPVFKDGGFTANDRDRVRYALRKTII 120
QY 120 RNIDLAELGAKYVAVWGREGAESGAADKDVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
DB 121 GNIDLAELGAKYVAVWGREGAESGAADKDVRAALDRMKEAFDILLGEYVTAQGYDLRFAI 180
QY 180 EKPNEPRGDILLPTTGHALAFIERLERPELYGVNPEVGEQWAGLNFPHGIAQALWAGK 239
DB 181 EKPNEPRGDILLPTTGHALAFIERLERPELYGVNPEVGEQWAGLNFPHGIAQALWAGK 240
QY 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGVW 299
DB 241 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGVW 300
QY 300 ASAAGCMRNYLILKERAAPRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDF 359
DB 301 ASAAGCMRNYLILKERAAPRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDF 360
QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
DB 361 DVDAARGMAFERLDQALMDHLLGARG 388

RESULT 5
US-07-640-476-11
Sequence 11, Application US/07640476
Patent No. 5376536

GENERAL INFORMATION:
APPLICANT: QUAX, WILHELMUS
APPLICANT: LUITEN, RUDOLF G.M.
APPLICANT: SCHUURHUIZEN, PAUL W.
APPLICANT: MRABET, NADIR
TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND
TITLE OF INVENTION: THEIR USE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/640,476

```

; FILING DATE: 19910110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kate H. Murashige
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20009.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Streptomyces violaceoniger
; STRAIN: CBS 409.73
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..389
; PUBLICATION INFORMATION:
; AUTHORS: Tiraby,
; JOURNAL: Nucleic Acids Res.
; VOLUME: 16
; PAGES: 9337-
; DATE: 1988
;
US-07-640-476-11

```

Query Match 93.3%; Score 1892.5; DB 1; Length 389;
 Best Local Similarity 92.3%; Pred. No. 1.5e-181;
 Matches 359; Conservative 15; Mismatches 12; Indels 3; Gaps 2;

```

QY 1 MNYQPTPEDRTFGLWTVGWRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 MSFQTPEDKFTFGLWTVGWRDPFGDTRPALDPVETVQRLAELGAYGVTFHDDDLIP 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 FGASDTEREAVKRPQALDAGTMTVPMTNLFTHPVFKAGFTANDRAVRRYALRKT 119
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 FGSSDTERESHKRPQALDAGTMTVPMTNLFTHPVFKDGGFTANDRAVRRYALRKT 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 120 RNIDLAVELGAKVYVWVGREGAESAADKVRALDRMKEAFDILLGEYVTSOGYDIRAI 179
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 RNIDLAAELGAKVYVWVGREGAESAADKVRALDRMKEAFDILLGEYVTAQGYDLRF 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 180 EPKPNPGRDILLPTIGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 239
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 181 EPKPNPGRDILLPTVGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 240 LFHIDLNGSGIKYDQDLRFAGAGDLRAAFWLVLDLESAGWEGPRHFDKPPRTEDIDGVW 299
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 241 LFHIDLNGSGIKYDQDLRFAGAGDLRAAFWLVLDLESAGYEGPRHFDKPPRTEDIDGVW 300
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 300 ASAGCMENYLLKERRAAFRADPEVQEARLRAARDLQAEPTAADGLQALLADRTAFEDF 359
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 301 ASAGCMENYLLKERRAAFRADPEVQEARLRAARDLQAEPTAADGLQALLADRTAFEDF 360
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 360 DVDAARGMA--FERLDQALDHLHGARG 386
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 361 DVEAARAARAFERLDQALDHLHGARG 389
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

RESULT 6
 US-07-637-870-3
 ; Sequence 3, Application US/07637870
 ; Patent No. 5310665
 ; GENERAL INFORMATION:
 ; APPLICANT: Lambier, Anne-Marie
 ; APPLICANT: Lasters, Ignace

```

; APPLICANT: Quax, Wilhemus J.
; APPLICANT: Van Der Laan, Jan M.
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES HAVING
; TITLE OF INVENTION: ALTERED SUBSTRATE SPECIFICITY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/637,870
; FILING DATE: 19910104
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20009.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-07-637-870-3

```

Query Match 91.6%; Score 1859; DB 1; Length 387;
 Best Local Similarity 91.2%; Pred. No. 3.4e-178;
 Matches 353; Conservative 15; Mismatches 17; Indels 2; Gaps 2;

```

QY 1 MNYQPTPEDRTFGLWTVGWRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 MNYQPTPEDRTFGLWTVGWRDPFGDTRPALDPVSRVRLSELGAYGVTFHDDDLIP 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 FGASDTEREAVKRPQALDAGTMTVPMTNLFTHPVFKAGFTANDRAVRRYALRKT 119
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 FGSSDTERESHKRPQALDAGTMTVPMTNLFTHPVFKDGGFTANDRAVRRYALRKT 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 120 RNIDLAVELGAKVYVWVGREGAESAADKVRALDRMKEAFDILLGEYVTSOGYDIRAI 179
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 121 RNIDLAVELGAKVYVWVGREGAESAADKVRALDRMKEAFDILLGEYVTSOGYDLRF 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 180 EPKPNPGRDILLPTIGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 239
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 181 EPKPNPGRDILLPTVGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 240 LFHIDLNGSGIKYDQDLRFAGAGDLRAAFWLVLDLESAGWEGPRHFDKPPRTEDIDGVW 299
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 241 LFHIDLNGSGIKYDQDLRFAGAGDLRAAFWLVLDLESAGYAGPRHFDKPPRTEDIDGVW 300
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 300 ASAGCMENYLLKERRAAFRADPEVQEARLRAARDLQAEPTAADGLQALLADRTAFEDF 359
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 301 ASAGCMENYLLKERRAAFRADPEVQEARLRAARDLQAEPTAADGLQALLADRTAFEDF 360
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 360 DVD-AAARGMAFERLDQALDHLHGARG 385
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 361 DVDAARGMAFEHLQALDHLHGARG 387
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

RESULT 7
 US-07-637-399-9


```

QY   360 DVD-AAARGMAFERLDOLANDHLLGAR 387
      ||| ||||| ||||| ||||| |||||
Db   361 DVDAARGMAFEHLDOQLANDHLLGAR 387

RESULT      8
US-07-640-476-12
; Sequence 12, Application US/07640476
; Patent No. 5376536
; GENERAL INFORMATION:
; APPLICANT: QUAX, WILHELMUS
; APPLICANT: LUITEN, RUDOLF G.M.
; APPLICANT: SCHUURHUZEN, PAUL W.
; APPLICANT: KRABET, NADIR
; TITLE OF INVENTION: NOVEL GLUCOSE I
; TITLE OF INVENTION: THEIR USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0.
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,476
; FILING DATE: 19910110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kate H. Murashige
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces violaceorub
; STRAIN: LMG 7183
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..387
; US-07-640-476-12

```

```

; LENGTH: 387 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces violaceoruber
; STRAIN: LMG 7183
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..387
; US-07-640-476-12

```

```

; ORIGINAL SOURCE:
; ORGANISM: Streptomyces violaceoruber
; STRAIN: LMG 7183
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..387
;
us-07-640-476-12

Query Match          91.6%;   Score 1859;   DB 1:   Length 387;
Best Local Similarity 91.2%;   Pred. No. 3.4e-478;
Matches 353;   Conservative 15;   Mismatches 17;   Indels 2;   Gaps 2;

QY      1  MNYOPTEDRFTFGLTWVGWGRDPFGDTPALDPVD-VORLAEAGYGVTFHDDDLIP 59
        1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1  MNYOPTEDRFTFGLTWVGWGRDPFGDTPQALDPAESVRLSELGAYGVTFHDDDLIP 60
        1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      60  FGASDTEREAHVKFRQALDATTGMTVPMATNTLTFHPVKAGATTANDRAYRRYALRKTI 119
        1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61  FGSDTERESHVKFRQALDATTGMKVPMTNTLTFHPVKDGAFTANDRDVRRYALRKTI 120
        1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      120  RNIDLAVELGAKYVAVWGREGAESGAAKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
        1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121  RNIDLAVELGASVYVAVWGREGAESGAAKDYRDALDRMKEAFDILLGEYVTEGGYDLKFAI 180
        1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 180 EPKPNRPGDILLPTIGHALAFTERLERPELYGVNPEVHQEAGLNFPHGIAQALWAGK 239
Db 181 EPKPNRPGDILLPTVGHALAFTERLERPELYGVNPEVHQEAGLNFPHGIAQALWAGK 240
QY 240 LFHIDLNGQSIIKYDDOLRFEGAGDLRAAFWLVDLLESAGMEGPRHDFDKPPRTEDIDGVW 299
Db 241 LFHIDLNGQSIIKYDDOLRFEGAGDLRAAFWLVDLLERAGYGRHDFDKPPRTEDFDGVW 300
QY 300 ASAAGCMRNYYLIILKERAAAFRADPVEQEAALRAARLDQLAEPTAADGLQALLADRDAYEDF 359
Db 301 ASAAGCMRNYYLIILKDRAAAFRADPQVEALAAARULDELARPTAEDGLAALLADRSAYDTF 360
QY 360 DVD-DAARGMAFERLDQLAMDHLLGAR 385
Db 361 DVDAAAARGMAFEHLDLQLAMDHLLGAR 387

RESULT 9
US-08-112-703-9
; Sequence 9, Application US/08112703
; Patent No. 5384257
; GENERAL INFORMATION:
; APPLICANT: Lambair, Anne-Marie
; APPLICANT: Lesters, Ignace
; APPLICANT: Mrabet, Nadir
; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Van Der Lean, Jan M.
; APPLICANT: Missel, Onno
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/112,703
; FILING DATE: 26-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 4615002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-112-703-9

```

```

Db      61 FGSSDTERESHIKRFRQALDATGKVPYMATNLFTHPVFKDGAFANDRVRYALRKTI 120
QY     120 RNTIDLAVELGAKVYVWAGREGAESGAADVRAALDRMKEAFDLLGEYVTSGYDIRFAI 179
Db     121 RNTIDLAVELGASVYVWAGREGAESGAADVRLDRMKEAFDLLGEYVTEOGYDLUKFAI 180
QY     180 EPKPNEPRGDILLPTIGHALAFIERLERPELYGVNPVEVGHEQMAGLNFPHGIAQALWACK 239
Db     181 EPKPNEPRGSDILLPTVGHALAFIERLERPELYGVNPVEVGHEQMAGLNFPHGIAQALWACK 240
QY     240 LFHIDLNGSQSIKYDDLRFGAGDLRAAFWLVDLLESAGMEGRHFDFPKPRTEDIDGYW 299
Db     241 LFHIDLNGSQSIKYDDLRFGAGDLRAAFWLVDLLERAGYAGRHFDFPKPRTEDIDGYW 300
QY     300 ASAAGCMRNYYLILKERAAAFRADPEVQEALRAARLDQLAEPTAADGLQALLADR'AYEDF 359
Db     301 ASAAGCMRNYYLILKDRAAFRADPPQVEALAAARLDLARPTAEDGLAALLADRSAYDTF 360
QY     360 DVD-AAARGMAFERLDQLAMDHLLGAR 385
Db     361 DVDAAAARGMAFEHLDDLQAMDHLLGAR 387

RESULT 10
US-07-637-399-8
; Sequence 8, Application US/07637399
; Patent No. 5340738
; GENERAL INFORMATION:
; APPLICANT: Lamber, Anne-Marie
; APPLICANT: Lasters, Ignace
; APPLICANT: Mrabet, Nadir
; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Van Der Lean, Jan M.
; APPLICANT: Misset, Onno
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN
; TITLE OF INVENTION: ALTERED PH OPTIMUM
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/637,399
; FILING DATE: 19910104
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 4615-0020.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; PS-07-637-399-8

```


TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-637-870-2

Query Match 66.7%; Score 1353.5; DB 1; Length 394;
Best Local Similarity 66.0%; Pred. No. 1.4e-127;
Matches 260; Conservative 46; Mismatches 79; Indels 9; Gaps 4;

QY 1 MNYQPTPEDRFTFGLTWGQGRDPFGDATRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
DB 1 MSQATPDDKFSFGLTWGQARDAFGDATRPVLDPIEAVHKLAEIGAYGVTFHDDDLVP 60
QY 60 FGADATREAHVKRFRQALDGMTVPMATTNLFTHPVFKAGAFANDRAVRRYALRKTI 119
DB 61 FGADATRGIVAGFSKALDETGLIIVPMVTNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120
QY 120 RNIDLAELGAKYVAVWGREGAESAADVRAALDRMKEAFDGLLGEYVTSQGYDIRFAI 179
DB 121 ROMDLGAELGAKTLVLWGREGAEDSAKDVGAAALDRYREALNLLAQYSEDQGYGLPFAI 180
QY 180 EKPNEPRGDILLPTTIGHALAFIERLERPELVGNPEVGEQMGAGLNFPHGIAQALWAGK 239
DB 181 EKPNEPRGDILLPTAGHAI AFVQELERPELVGINFETGHEQMSNLFNTQGIQAALWHKK 240
QY 240 LFHIDLNGSGIKYDQDLRFAGDLRAAFWLVDLLESA-----GWEGPRHFDKPPRTED 294
DB 241 LFHIDLNGHGHKFDQDLVFGHGDLLNAFSLVDLLENGPDGPGPAYDGPFRHFDKPPRTED 300
QY 295 IDGVWASAGCMRNLYILKERAAAFRADPEVQALRAALDQAEPTAADG--LQALLAD 352
DB 301 FDGVWESAKDNIRMYLLKRAKAFRADPEVQAALAESKVDLRTPTLNPGETYADLLAD 360
QY 353 RTAYEDFDVDA-AARGMAFERLDOLAMDHLGLAR 385
DB 361 RSAFEDYDADAVGAKGYGVFKLNQLAIDHLLGAR 394

RESULT 15

US-07-637-399-2
Sequence 2, Application US/07637399
Patent No. 5340738

GENERAL INFORMATION:
APPLICANT: Lambelr, Anne-Marie
APPLICANT: Lasters, Ignace
APPLICANT: Mrabet, Nadir
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Van Der Laan, Jan M.
APPLICANT: Misset, Onno
TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN
TITLE OF INVENTION: ALTERED PH OPTIMUM
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Morrison & Foerster
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/637,399
FILING DATE: 19910104
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 4615-0020.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-637-399-2

Query Match 66.7%; Score 1353.5; DB 1; Length 394;
Best Local Similarity 66.0%; Pred. No. 1.4e-127;
Matches 260; Conservative 46; Mismatches 79; Indels 9; Gaps 4;

QY 1 MNYQPTPEDRFTFGLTWGQGRDPFGDATRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
DB 1 MSQATPDDKFSFGLTWGQARDAFGDATRPVLDPIEAVHKLAEIGAYGVTFHDDDLVP 60
QY 60 FGADATREAHVKRFRQALDGMTVPMATTNLFTHPVFKAGAFANDRAVRRYALRKTI 119
DB 61 FGADATRGIVAGFSKALDETGLIIVPMVTNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120
QY 120 RNIDLAELGAKYVAVWGREGAESAADVRAALDRMKEAFDGLLGEYVTSQGYDIRFAI 179
DB 121 ROMDLGAELGAKTLVLWGREGAEDSAKDVGAAALDRYREALNLLAQYSEDQGYGLPFAI 180
QY 180 EKPNEPRGDILLPTTIGHALAFIERLERPELVGNPEVGEQMGAGLNFPHGIAQALWAGK 239
DB 181 EKPNEPRGDILLPTAGHAI AFVQELERPELVGINFETGHEQMSNLFNTQGIQAALWHKK 240
QY 240 LFHIDLNGSGIKYDQDLRFAGDLRAAFWLVDLLESA-----GWEGPRHFDKPPRTED 294
DB 241 LFHIDLNGHGHKFDQDLVFGHGDLLNAFSLVDLLENGPDGPGPAYDGPFRHFDKPPRTED 300
QY 295 IDGVWASAGCMRNLYILKERAAAFRADPEVQALRAALDQAEPTAADG--LQALLAD 352
DB 301 FDGVWESAKDNIRMYLLKRAKAFRADPEVQAALAESKVDLRTPTLNPGETYADLLAD 360
QY 353 RTAYEDFDVDA-AARGMAFERLDOLAMDHLGLAR 385
DB 361 RSAFEDYDADAVGAKGYGVFKLNQLAIDHLLGAR 394

Search completed: March 22, 2000; 23:43:55
Job time: 781 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2000, 23:04:53 ; Search time 752.35 Seconds
(without alignments)
-6239.449 Million cell updates/sec

Title: US-09-383-318-1
Perfect score: 1546
Sequence: 1 cagagcgcttggtgact.....gcgcggtgtgggcgcgtgc 1546

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 821193 seqs, -1518192014 residues
Total number of hits satisfying chosen parameters: 1642386

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Result No.	Score	Query Match	Length	DB ID	Description
1	1546	100.0	1546	1	SSPY15518
2	1031.4	66.7	1952	2	AF170068 Streptomyces
3	977.2	63.2	3060	1	STMXYLAB
4	972.6	62.9	1201	1	STMXYLI
5	971	62.8	1202	1	SVXYL
6	970.8	62.8	1522	1	S73809
7	952.8	61.6	1164	5	AL0243 S.murinus D
8	947.8	61.3	1164	5	E15670 Streptomyces
9	935	60.5	1167	5	I09223 Sequence 1
10	904.6	58.5	2010	5	E01992 DNA sequenc
11	625.2	40.4	1892	1	AMSXYLI
12	599.6	38.8	1567	5	A06406 Nucleotide
13	599.6	38.8	1639	1	AMXI
14	598.8	38.7	1185	5	A06404 Nucleotide
15	598.4	38.7	1182	5	AL0241 A.missourie
16	531.4	34.4	1905	1	ASXYLA
17	481.8	31.2	1805	1	THXYLA
18	481.8	31.2	1805	5	I11743 Sequence 1
19	108.6	7.0	41622	2	SCD25
20	103	6.7	37750	2	SC6D7
21	99.2	6.4	42655	1	SC7H2
22	98.6	6.4	1820	1	STMSLPD
23	98.6	6.4	40544	1	SC5
24	98.6	6.3	38404	1	SC2G5
25	98	6.3	1821	5	AR027260
26	97	6.3	1821	5	AR027260
27	96.8	6.3	182526	40	AC007731
28	96.8	6.3	192592	40	AC005500
29	96.2	6.2	24700	1	SCB8
30	95.2	6.2	150724	10	AP000550
31	95	6.1	3925	2	AF177946
32	95	6.1	23990	1	SC6E10
33	94.4	6.1	7412	16	PVUL50S
34	94	6.1	19830	1	SC3F9
35	93	6.0	5530	1	STMFP5
36	93	6.0	25306	2	AF058302
37	92.2	6.0	22449	1	SPSNBCDE
38	92.2	6.0	22449	1	SPSNBCGEN
39	91.2	5.9	22115	1	SCC22
40	91.2	5.9	40909	1	SCI35
41	89.8	5.8	23730	1	SC3F7
42	89.4	5.8	45396	1	SCH35
43	89.4	5.8	1593	2	AF113605
44	89.2	5.8	5800	1	STMCHO
45	88.8	5.7	7836	2	AF047717

ALIGNMENTS

RESULT 1
SSPY15518 1546 bp DNA
LOCUS Streptomycetes sp. SK strain xyla gene.
DEFINITION Y15518
ACCESSION Y15518
VERSION Y15518.1 GI:4210848
KEYWORDS xyla gene.

SSPY15518 1546 bp DNA
Streptomycetes sp. SK strain xyla gene.
Y15518
Y15518.1 GI:4210848
xyla gene.

01-FEB-1999

|||||
Db 1501 GGGCCCATGCTGCTGCTCCCGGGCGCGGTGTGGGGCGGTGC 1546

RESULT 2

AF170068
LOCUS AF170068 1952 bp DNA BCT 12-AUG-1999
DEFINITION Streptomyces chibaensis D-xylose isomerase (xylA) gene, complete cds.
ACCESSION AF170068
VERSION AF170068.1 GI:5731284
KEYWORDS Streptomyces chibaensis.
SOURCE Streptomyces chibaensis.
ORGANISM Streptomyces chibaensis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 1952)
Joo, G.J., Shin, J.H., Heo, G.Y., Park, H.D. and Rhee, I.K.
Streptomyces chibaensis J-59 xylA
Unpublished
2 (bases 1 to 1952)
Joo, G.J., Shin, J.H., Heo, G.Y., Park, H.D. and Rhee, I.K.
Direct Submission
Submitted (19-JUL-1999) Department of Agricultural Chemistry,
College of Agriculture, Kyungpook National University, Buk-gu,
Sankyuk-Dong, 1372, Taegu 702-701, Korea

FEATURES

source
1..1952
/organism="Streptomyces chibaensis"
/strain="J-59"
/db_xref="taxon:67286"
552..1718
/gene="xylA"
552..1718
/gene="xylA"
552..1718
/codon_start=1
/transl_table=11
/product="D-xylose isomerase"
/protein_id="A484850.1"
/translation="MSYQPTPEDRFETFLGTVWGQRPFGDTRALDPVETVQRLA
ELGAGVTHDDLLPFSSDPTRESHIKRFQALDAGMTVPMTATNLFTHPVKDG
AFTANDRVRYALAKTIRNIDLAELGAKTVVANGREGAESGAKDVRSLDRMKE
AFDLDGVTSQGLRFAIEKPNRPGDILLPVGHAFIERLEPXYNPEV
GHEQAGLVPHGIAQALWAGKFIHDLNGSGIKYDQDLRFAGDLRSFVLDLLE
SAGYEGPHFDPKPTRELDGWSAAGCMRNLYLKERAAAFRAFDPEVQAALRAS
LDLQAQTAADGLELDLADRAFEFDVEAAAARGMAFERLDQALMDHLGARG"

gene

552..1718

CDS

552..1718

BASE COUNT 286 a 740 c 654 g 272 t
ORIGIN
Query Match 66.7%; Score 1031.4; DB 2; Length 1952;
Best Local Similarity 85.0%; Pred. No. 1.6e-84;
Matches 1228; Conservative 0; Mismatches 176; Indels 41; Gaps 5;
QY 2 acgagcgcttggtgactggtgagcgagtcacacacgagcagcgagcgacccctcggc 61
|||||
Db 326 ACAGAGCGCTTGTGACTGGGTGGACGTGTCCACTCCGACGACGA-GCGACCCCTCGC 384
QY 62 tgcgacatcgactccctctctttcccggtcagggactgacactgaggttcacgc 121
|||||
Db 385 TCGTGACATCGGGCTCTCCCTTCGCGGGCTCCGCGGACGTGCGGTGTGTGTTTCA 444
QY 122 tatgcggcgctgtggcgcccggtgtgcgacccgccccggcgctttctgttcgcgt 181
|||||
Db 445 AGACACCTTGTCTCGAAGACATCTTG----- 472
QY 182 tccctcccgagcgctgcggcactactaattgttaatcgccctgacgaaatagtcgc 241
|||||
Db 472 TCTCTTCCAGAGGTGCG-TCGGCATACTAATTTGTAACCGGCATGACGAATAGTCGC 530
QY 242 asgagcagcagcgcggcgtgaactaccagccaccccgagcagaggttcaccttc 301
|||||

Qy	1323	ccgcgtacgaggacttcagacgttgacgcggc-----cgccgcgcgcgatggccc-ttcgaagc	1376
Db	1087	CCGCGTTTCGAGGACTTCGACGTGTGAGGGCGGCCGCGGCATGCGCGTTCGAAC	1146
Qy	1377	gcctcgaccagctgcgcattgacgacacacctgtggcgccgcggcgctgaacggcg	1431
Db	1147	GCCTCGACCAGCTGGCGATGACCACTGCTGGCGCGCGCGGCTGACGGTCCGC	1201
 RESULT 5			
SVXYL 1202 bp DNA BCT 18-JUL-1995			
LOCUS Streptomyces violaceoniger xyIA gene for D-xylose-isomerase			
tetramer alpha 4 ; EC 5.3.1.5).			
X12816			
ACCESSION X12816.1 GI:48004			
VERSION xyIA gene; xylose isomerase.			
KEYWORDS Streptomyces violaceoniger.			
SOURCE Streptomyces violaceoniger.			
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.			
REFERENCE 1 (bases 1 to 1202)			
AUTHORS Drocourt,D.			
TITLE Direct Submission			
JOURNAL Submitted (09-SEP-1988) Drocourt D., CAYLA, Avenue de Larrieu,			
31094 Toulouse Cedex, France			
REFERENCE 2 (bases 1 to 1202)			
AUTHORS Drocourt,D., Bejar,S., Calmels,T., Reynes,J.P. and Tiraby,G.			
TITLE Nucleotide sequence of the xylose isomerase gene from Streptomyces			
violaconiger			
JOURNAL Nucleic Acids Res. 16 (19), 9337 (1988)			
MEDLINE 89016633			
FEATURES Location/Qualifiers			
1..1202			
/organism="Streptomyces violaceoniger"			
/strain="CBS 409-73"			
/db_xref="taxon:1953"			
12..17			
/note="pot. ribosome binding site"			
RBS 25..1194			
CDS /note="D-xylose isomerase (AA 1 - 389)"			
/codon_start=1			
/transl_table=11			
/protein_id="CAA31304.1"			
/db_xref="GI:48005"			
/db_xref="SWISS-PROT:P09033"			
/translation="MSFQPTPEKFTFLMTWQGRDPGDTPALDVEYVQLA			
ELGAVTFHDDLLIPFGSSDTRESHIKFERQALDATGTPMVAINTLFTHPVFKDG			
GFTANDRVRYALRTINRIDLAELGAKTYVAMGREAESGGAKDVDRDALDRNKE			
AFDLQEVTAQFDUNFALEPNPREDGILLIPTGHALAFIERLERPELYGVNPVEV			
PHQMGAGLNPHQIQAQWAGKLFIHDNLNGSQIKYDQURFAGLDRAAFWLVDLLE			
SDGYEPRHFDFPFPTEDPDGWASAEQCMRNILILKERAAFRAADPEVQEALRAAR			
LQLOAPTDAGTEALLADRTADEDFDEAAAAAAWPPERLDQLAMDHLGLRG"			
BASE COUNT 187 a 454 c 403 g 158 t			
ORIGIN			
 Query Match 62.8%; Score 971; DB 1; Length 1202;			
Best Local Similarity 90.0%; Pred. No. 4.7e-79;			
Matches 1076; Conservative 0; Mismatches 110; Indels 9; Gaps 3;			
Qy	246	gagcaagagcgcgcatgaactaacagccccccccccagagcacggttaccttcggcc	305
Db	8	GAAACAAGGAGCAGCGCATGAGCTTCAGGCCACCCCGGAGGACAAGTTCACCTTCGGTC	67
Qy	306	tgtgacgcgtcgctgcagggcgcgacccttcgccgacgcacacgcgtcccccctcg	365
Db	68	TGTGACCGCTCGGCTGGCAGGGAAGGACCGCTTCGGCGAGCCACCGCCCTGCCCTCG	127
Qy	366	accgggttgcg---cgtgcagcggtgcgcgaactggggcgctcagcgagtgcattccacg	422
Db	128	ACCGCGTCGAGACCGTGACGCGCTGGCCGAGCTGGGGCCCTACGGAGTAGCTTCACG	187

FT /organism='Streptomyces griseofuscus' FT CDS
1. .1164 /product='glucose isomerase'.
Location/Qualifiers
1. .1164 /organism='Streptomyces rochei'
/db_xref='taxon:1928' 155 t
BASE COUNT 181 a 441 c 387 g
ORIGIN
Query Match 61.3%; Score 947.8; DB 5; Length 1164;
Best Local Similarity 90.1%; Pred. No. 5.5e-77;
Matches 1051; Conservative 0; Mismatches 107; Indels 9; Gaps 3;
QY 263 atgaactaccagcccccagggagagaggttcaccttcggcctggtgagcgtcggtg 322
DB 1 ATGAGCTTCAGCCACCCCGAGGACAGTTACCTTCGGTCTGTGGACCGTCGGCTG 60
QY 323 caggggcgagacccttcggcgagcccaagcgctccgcgcctcaccoggtcg 379
DB 61 CAGGGAAGGAGACCGTTGCGGAGCGGACACCCGCCCTGGCTCGACCGGTCGAGACCGTG 120
QY 380 cagcgctggcggaactggcgctcagagtgacgtacacccacagacagcctgacccc 439
DB 121 CAGCGCTGGCGAGGTGGCGGCTACGGAGTGACCTTCACGACGACGACGATGCC 180
QY 440 ttccggggtcccgacacccagcgagcgacgtcgaagcggttcgcgtcagcgctcag 499
DB 181 TTCGGGTCTCCACACCGAGCGGAGTCGCATACAGCGGTTCGCGACGGCCCTGGAC 240
QY 500 ggcagggcgatgaccttcgcgttcgagcccaacacacacacacacacacacacac 559
DB 241 GCCACTTGGCATGACGCTGCGATGGCCACCAAGCAACCTCTACCCACACCCCGTTCAG 300
QY 560 gcaggcggttcacccgacacccagcgagcgacgtcgcgcgttcacccctcgcaagacc 619
DB 301 ---GACCGTTTACCGCCCAACGACCGGACGTCGGCGCTACCGCGTGCACAGACGATC 357
QY 620 cggacacatcgatcgcggtcgagctggcgcaaggtcagctcgctggcgcgccgc 679
DB 358 CGCAACATCGACTGCGGCGGAGCTGGGCGCCAGACGTACGTGCGCTGGCGGCGCGT 417
QY 680 gaggcgcgagtcggtgcgcgaagagcgctgctgctggcgccctggacccatgaaggag 739
DB 418 GAGGCGCGGAGTCCGCTGGCGGCAAGGACGTGCGGACGCGCTCGACCGCATGAAGGAG 477
QY 740 gcttcgacctgctggcgagtagtcacacacacacacacacacacacacacacacac 799
DB 478 GCGTTGACCTTCGCGGAGTACGTACCGCGCCAGGGCTACGACCTCCGCTTCGCCATC 537
QY 800 gaggcccaagcgaacagcgcgcgagacatcgtgctccacacacacacacacacacac 859
DB 538 GAGCCCAAGCCCAAGGACCCCGGCGGACATCTCTGCGCCACCGCTCGCGCACCGCCTG 597
QY 860 gccttcacagcgctggagcgcccgagctgacggtgtcaaccccgaggtggggccac 919
DB 598 GCCTTCATCAGCGCTTGAGCGCCCGAGCTGTACGGCGTCAACCGGAGGTTCGGCCAC 657
QY 920 gacagatggcgcgctgaacttcgcgcacggaatcgatcgctgctggcgcgcgctg 979
DB 658 GAGCAGATGGCGCGCTGAATCTCCCGACGCGATCGCGAGGCGCTGTGGGCGGGCAAG 717
QY 980 ctcttcacatcgacctcaagcgccagtcggcgtacgaatcagacacacacacacacac 1039
DB 718 CTCTTCACATCAGCTCAACGCGCCAGTCCGGCATCAAGTACGACCAAGGACCTTCGGCTTC 777
QY 1040 ggcgcgggtgaactgcgcgcgcttcctgctgctgacacacacacacacacacacacac 1099
DB 778 GGGCGCGGCGACCTGCGGCGCGGCTTCGCTGCTGCTGACCTCTCTGGAGAGCGCGGTAC 837
QY 1100 gagggtccgcgcaccttcgacttcgaagcccccgcgagaccgagacatcgacgcggtg 1159

DB 838 GAGGGCCCGCGGACACTTCGACTTCAAGCGCGCGGACCGGAGGACTTCGACGCGGTGTG 897
QY 1160 gctcccgcgcggtgcatgagcgaactacacacacacacacacacacacacacacacac 1219
DB 898 GCCTCGCGGAGGGCTGCATGCGCAACTACCTGATCTCAAGGAGCGCGGCGCGCTTC 957
QY 1220 cgtgcgaccccgaggtccagagggccctcgctgctgcgcgcgcgcgcgcgcgcgcgc 1279
DB 958 CGCGCGGACCGGAGGTGAGGAGGCGCTGCGCGCGCGCTGACACAGCTGGCCAG 1017
QY 1280 ccacacgcgcgacgctcagcgccctgctgctgacacacacacacacacacacacacac 1339
DB 1018 CCACCGCGCGGACGCGCTGAGGCGCGCTGCTCGCGACCGACCGCTTCGAGGACTTC 1077
QY 1340 gacgtggaacgcg---gcccgcgcgcgctgagcgcttcgagcgcttcgacacacacacac 1396
DB 1078 GACGTGGAGGCGCGCGCGCGCGCATGCGTTCGAACGCGCTCGACGACGCTGGCGATG 1137
QY 1397 gaccacactgctggcgcgcgcgcgcgctga 1423
DB 1138 GACCACCTGCTGGCGCGCGCGCTGA 1164
RESULT 9
LOCUS I09223 1167 bp PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8901520.
ACCESSION I09223
VERSION I09223.1 GI:588068
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1167)
AUTHORS Drummond, R.J., Bloch, W., Matthews, B.W., Toy, P.L. and Nicholson, H.H.
TITLE PROCARYOTIC XULOSE ISOMERASE MUTAINS AND METHOD TO INCREASE PROTEIN STABILITY
JOURNAL Patent: WO 8901520-A 1 23-FEB-1989;
FEATURES
Location/Qualifiers
Source 1. .1167 /organism='unknown'
BASE COUNT 180 a 426 c 405 g 156 t
ORIGIN
Query Match 60.5%; Score 935; DB 5; Length 1167;
Best Local Similarity 88.8%; Pred. No. 7.6e-76;
Matches 1036; Conservative 0; Mismatches 125; Indels 6; Gaps 2;
QY 263 atgaactaccagcccccagggagagaggttcacacacacacacacacacacacacacacac 322
DB 1 ATCACTTACAGCCACCCCGGAGGACAGGTTTACCTTCGACTGCGACCGTCGGCTGG 60
QY 323 caggcgcgagaccccttcggcgagcgacgctcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 379
DB 61 CAGGAGCGGAGACCGCTTCGCTGACCGCGCGCGCGCTCGACCGGCTCGAGTCCGTG 120
QY 380 cagcgctggcggaactggcgcgctcagagtgacacacacacacacacacacacacacacac 439
DB 121 CGCGCGCTGCGGAGCTGGGCGCGCGCGCGCGCGCGCTCGCTTCACGACGACGACCTCATCCCC 180
QY 440 ttccggggtcccgacacccagcgagcgcgaggttcacggttcacacacacacacacacacacac 499
DB 181 TTCGGTCCAGCGACGACGAGCGGAGGACGACGTCAAGCGGTTCGGGACGGCGCTGGAC 240
QY 500 ggcacggcgatgaccttcgcgttcgagtcgagcccaacacacacacacacacacacacacac 559
DB 241 GACACCGCATGAAGTGGCGATGGCCACCAACCTGTTACCCACCGCGGTTCAG 300
QY 560 gcaggcgcttcacccgcaac 619
DB 301 GACGGGCGGCTTACCGCCCAACGACCGCGGCTACGCGCGCTGCGCCCTGCGCAAGACCATC 360


```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 999 GCTACGCGTGGCCAGAGATCCCAACATCGACTGGCGCCGAGTGTCGGCGCAAG 1058
QY 656 gtctagctgcctgggcccgcgagggcgagtcctccggtgcccgaagagcgtgcgt 715
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1059 ACGTAGCTCGCTGGCGCGCGCTGAGGCGCGGAGTCCGGTGGCGCCAGGAGCTGCGC 1118
QY 716 gggccctggaccgcatgaagggagccttcagacctcctgcgcgagtcacgtcaacctgcag 775
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1119 GAGCGCCTCGACCGCATGAAGGAGGGTTCTGACCTCTCTCGCGGAGTACGTCACCGCCAG 1178
QY 776 ggtctagacatcgggttcctccatcagcccaagcccaagcgcgcgcgcgcacatcctg 835
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1179 GGTACGACCTCGCTTCGCCATCGAGCCCAACGCCCAACAGAGCCCGCGCGGACATCCTC 1238
QY 836 ctgcccaccatcgccacgcgctcgcttcctcatcgagcgctggagcgcccccgagctgtac 895
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1239 CTGCCACCGTCGGCCACGCCCTTTCATCATGAGCGCTGGAGCGCGCGGAGCTGTAC 1298
QY 896 ggtgtcaacccccagatgggcccacagacagatggcgcgccctgaactcccgacgcgcac 955
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1299 GCGCTCAACCCGAGGTGCGCCACGACGACAGATGCCGCGCTGAACCTCCCGCACGCGATC 1358
QY 956 ggcaggtctgtggcggaagctcttcacatcgacctcaacgcgcagtcgcgcac 1015
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1359 GCGCAGCCCTGTGGCGGCGCAAGCTTTCACATCGACCTCAACGGCCAGTCCGGCATC 1418
QY 1016 aagtagacacgagacctgcgcttcggcgccgctgacctgcgcgcgcgccttcgtgctgc 1075
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1419 AAGTAGACACGAGACTGCGG - - - TCGCGGCGCGCACCTGCGGCGCGGCTTCTGG - TGGTC 1475
QY 1076 gaactctggagagcgccgctgggaggttcgcgcgcacacttcgaactccaagccccgcgcg 1135
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1476 GACCTCTCTGAGAGCGCGGTTACGAGGCGCGCGCACCTTCGACTCAAGCGCGCGCGG 1535
QY 1136 accgagacatcagcggcgtgtggcctcgcgcgcgcgcgggtgcgtgcgcacactacctgac 1195
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1536 ACCGAGGACTTCACGCGCTGTGGCGCTCGCGCGAGGGCTGCATGCGCACTACCTGATC 1595
QY 1196 ctgaagagcgccgcgcgccttcctgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1252
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1596 CTCAAGCAGC - CGCGCGCGCTTCCGCGCGCGCACCCGAGGTGCAGGCGCGCGCTCGGC 1654
QY 1253 gcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1311
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1655 GCGCGCGCTGTGACGAGCTGGCCACGCGACCGCTGGCGGAGCGGCTGGAGCGCCCTGCT 1714
QY 1312 ggcgaccgacccgctacgagacttcgactgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1370
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1715 CGCGGACCGACCGCTTCAGGACTTCGAGCTGAGGCGCGCGCGCGCGCGCGCATGGT 1774
QY 1371 ttagcgcttcgaccagcttcgcatggaccacactgtctggcgcccgcgctgaacccggcg 1430
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1775 TCGAAGCGCTTCGACCAAGCTGGCGATGGACCACTGCTGGCGCGCGCGCGCTGACGCTGC 1834
QY 1431 gacgaggggtagcgcgcgtcgtatcctcctgctcgtcatgagggggtgctgaggcgctc 1490
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1835 GCCAGGAAGCGGTGACGGAAGAACCCGGGGGACGCGCGCGCTGCGCGCGCGCTGCTCC 1894
QY 1491 gaggcgcccgccg 1503
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1895 CGGCGCTTCGC 1907
```

RESULT 11

```

AMXYLI 1892 bp DNA BCT 21-MAR-1996
LOCUS Actinoplanes sp. ATCC 31351 xylose isomerase gene, complete cds;
DEFINITION putative xylose kinase gene, partial cds.
ACCESSION M15050
VERSION M15050.1 GI:141993
KEYWORDS xylose isomerase; xylose kinase.
SOURCE Actinoplanes ATCC31351 DNA.
```

ORGANISM

Actinoplanes ATCC31351
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Micromonosporineae; Micromonosporaceae;
Actinoplanes.

REFERENCE

1 (bases 1 to 1892)

AUTHORS

Saari,G.C., Kumar,A.A., Kawasaki,G.H., Insley,M.Y. and O'Hara,P.J.

TITLE

Sequence of the Ampullariella sp. strain 3876 gene coding for

JOURNAL

xylose isomerase

MEDLINE

J. Bacteriol. 169 (2), 612-618 (1987)

FEATURES

87109047

source

Location/Qualifiers

CDS

1..1892

FEATURES

/organism="Actinoplanes ATCC31351"

FEATURES

/strain="Dow Chemical Co. 3876"

FEATURES

/db_xref="taxon:1872"

FEATURES

220..1404

FEATURES

/codon_start=1

FEATURES

/transl_table=11

FEATURES

/product="xylose isomerase"

FEATURES

/protein_id="AA092578.1"

FEATURES

/db_xref="GI:141994"

FEATURES

/translation="MSLQATPDKFSFGLTWVGWQARDAFGDATRPVLDPIBAVHKLA

FEATURES

ETGAYCTFHDDLVFPGADATRDGIVAGFSKALDETLGIVPMTNLTFFHVFKDG

FEATURES

GFTSNDRSVRRYAIRKVLQMDLGAELGAKTLVLMGREGAIDYDAKQVGAALDRYRE

FEATURES

ALNLAQYSEDQGYGLPFAIEPKNEPRGDIPLTPAGHAIAFVQELERPELGIIDRET

FEATURES

GHEQSNLNETQGIQAALWHKLFHIDLNGOHGPKEDODLVFGHDLNFAISIVDLLE

FEATURES

NGPDGPDYDGRPHFDYRPSRTEDFDGVWESAKNIRMYLLIKERAKAFRADEVOAA

FEATURES

LAESKVDLELTPTNPGETYADLLADRSAFEDYDADAVGAKGYGFVKLNQLAIDHLGL

FEATURES

AR"

FEATURES

1407..>1892

FEATURES

/note="putative"

FEATURES

/codon_start=1

FEATURES

/transl_table=11

FEATURES

/product="xylose kinase"

FEATURES

/protein_id="AA092579.1"

FEATURES

/db_xref="GI:141995"

FEATURES

/translation="NALVAGIGQLDAVVGQHSATPRPATGPAGAAHSGRHOYDPDA

FEATURES

HWARTGDSREAGLGRLAAASVAGQQRHGALLESAVTVPRLALLNDTRPGAAA

FEATURES

DLIQELGGADKWAENAVGIVPVASLILTNSGLARHEPANAANKVAICLPHDMLTWKLS

FEATURES

GS"

FEATURES

BASE COUNT 319 a 671 c 643 g 259 t

FEATURES

ORIGIN 1033 bp upstream of the Sali site.

FEATURES

Query Match 40.4%; Score 625.2; DB 1; Length 1892;

FEATURES

Best Local Similarity 73.3%; Pred. No. 2.4e-48;

FEATURES

Matches 865; Conservative 0; Mismatches 288; Indels 27; Gaps 4;

FEATURES

QY 271 ccaagccaccccgagagcaggttcacctcgcctcgtgacccgtcggtcgagggcg 330

FEATURES

Db 228 CCAGGCCACACCCGATGACAAAGTTCCTTCGTCTCTGGACCGTCGCTGGCAGCGCG 287

FEATURES

QY 331 ggaccttcggcgacgcacgc 387

FEATURES

Db 288 TGACGCGTTCGGTGACGCCACCGTCCGGTCTCGACCCGATCGAGCCGCTGCACAGCT 347

FEATURES

QY 388 ggcgaactggcgctacgagtgaccttcacgacgacgacgacgacgacgacgacgacgac 447

FEATURES

Db 348 GCGCGAGATCGCGCGGTACGCGCTCACGTTCCACGACGACGACGACGACGACGACGAC 407

FEATURES

QY 448 gtccgacacgc 507

FEATURES

Db 408 CGACGCGCGACCGCGACGCGCATCGTCGCCGGGTCTCCAAAGCGCTCCACGAGACCGG 467

FEATURES

QY 508 catgacggttcgatggccaccaccacaccttcaccaccaccaccttcaccaccaccaccttc 567

FEATURES

Db 468 CCGTAGCTGCTCCGATGGTCACCAACCTGTTCAACCCACCGGTGTTCAAGGACGCGCG 527

FEATURES

QY 568 gttcacgcgaacgacgc 627

FEATURES

Db 528 CTTACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 587

FEATURES

QY 628 cgatctcgcggtcgagctggcgcccaaggtctacgtcgctggcgccgctggcgccgagggcg 687

Db 588 GGACCTCGCGCCGAGCTGGCGCCAGAACCTGTGCTCTGGGCGCGCCGAGGGCGC 647
 Qy 688 ggaagtcggtccgccaagagcggtgctgctgagccgcatgaagagggccttga 747
 Db 648 CGAGTACGACTCGGCCAAGGACGTCGGCGCCGCTCCTACCGCTACCGCGAGGCGCTCAA 707
 Qy 748 cctgctcgcgagtaagtcacactcgcagggctacgaacatcgggttcgcacatcgagcccaa 807
 Db 708 CCGTCTCGCGAGTACTCCGAGGACGAGGCTACGGCTGCGCTGCGCATCGAGCCGAA 767
 Qy 808 gccgaacgagccgagcgagcactcctgctgcccacacatcggccacgagcgtcgtcctcat 867
 Db 768 GCCGAACGAGCCCGCGCGGACATCCTGCTCCGAGCCGCGCCACGACATCGCGTTGCT 827
 Qy 868 cgagcgcctgagggcccgccgagcgtgaggtgtaaaccccgaggtggccacgagcagat 927
 Db 828 CGAGGAGTGGAGCGCCCGGAGCTTTCGGCATCAACCGGGAGACCGGCCACGAGCAGAT 887
 Qy 928 ggcggcctgaaactcccgagcgagcgcaggtcgtggtggggggaagcctttcca 987
 Db 888 CTCGAACCTGAATTCACCGAGGACATCGCCAGGCGTGTGGCACAAAGAGCTTCTCA 947
 Qy 988 catgacactcaacggccgagtcgagtcgaatgaagtaagacagacactgcttcggcgccgg 1047
 Db 948 CATGACCTGAACGCCACGACGACGCGCCGAAAGTTCACGAGGACCTGTCTTCGGTTCACGG 1007
 Qy 1048 tgaactcgcgccgcttctgctggttcgagcactgctgaga-----gcgc 1092
 Db 1008 TGACTGTCTCAACGGGTCTCCCTGGTGCAGCTCTTTGGAGAACGGGCCCGAGCGGCC 1067
 Qy 1093 cggctggaggttcgcccacttcgaactcaagcccgccgagccggagacatcgaagg 1152
 Db 1068 GGCCTACGACGCGCGCGGCACTTCGACTACAAAGCCCTCGCGACCGAGGACTTCGACGG 1127
 Qy 1153 cgtgtggcctccgagcgccgaggtgcatgcgaactacatcctgagggagcgccgc 1212
 Db 1128 CGTCTGGAGTCGGCCAAAGGACAAATCCGATGTACCTGCTGTCTCAAGAGCGGGCCAA 1187
 Qy 1213 cgccttcgctccgacccgaggttcagagagccctgctgctccgagcctcgaagcct 1272
 Db 1188 GCGGTTCGGGCGCACCGGAGGTGACGGCGGCTGCGCCGAGTCAAGGTCGACGAGCT 1247
 Qy 1273 -----cgccgagccacccgagcgccgagcgctgagggcctgctgctccgagcagccgc 1326
 Db 1248 CGGGACCCCGAGCTGAACCGCGGCGAGACCTACCCGACCTGCTGGCGACCGTAGCGC 1307
 Qy 1327 gtacgaggaactcagctgagcgagcgccgagcgcgc---ggcatggccttcgagcgccctga 1383
 Db 1308 GTTCGAGGACTACGACGCGGACGCGGTGCGGGCCAAAGGGCTACGGGCTTCGTCAAGCTCAA 1367
 Qy 1384 cgaactgcatggaacacactgctgagcgccgagcgctga 1423
 Db 1368 CCAGCTGGCGATCGACCACTGCTCGGAGCGCGGTGATCA 1407

RESULT 12

A06406 A06406 1567 bp DNA PAT 18-JUN-1993
 LOCUS Nucleotide sequence of the P(R)-gi transcriptional unit.
 DEFINITION A06406
 ACCESSION A06406
 VERSION A06406.1 GI:412855
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 1567)
 JOURNAL
 AUTHORS Patent: WO 900605-A 14 25-JAN-1990;
 FEATURES Location/Qualifiers
 1..1567 /organism="synthetic construct"
 /db_xref="taxon:32630"

BASE COUNT 300 a 532 c 490 g 245 t
 ORIGIN
 Query Match 38.8%; Score 599.6; DB 5; Length 1567;
 Best Local Similarity 71.0%; Pred. No. 5e-46;
 Matches 861; Conservative 0; Mismatches 324; Indels 27; Gaps 4;
 Qy 271 ccagccacccccagagacaggttcaactcgcgctctgagccctggtgagcaggggagc 330
 Db 197 CCAGCCACACGCGAAGAAAGTTCTCTTCGGTCTCTGGACCGTTGGATGGCAGGTCTG 256
 Qy 331 ggaaccccttggcagcgacgcagcgtcccgccctcgacccggtcga---cgtgcagggct 387
 Db 257 TGACGCGTTCGGTGACGCCACGCGTACGGCATCGAGCCGCTCGAGCCGCTGCAACAAGCT 316
 Qy 388 ggcgaactgggcccctacggagtgaccttcacagacagacactgatacccttcggggc 447
 Db 317 CGCTGAGATCGGCGCTACGGCATCAGCTTCCAGGACGACACCTGTGCTTCGGCTC 376
 Qy 448 gtcgacacccagcgagcgagcgacgtcaagcgttcctcagcgctcgacgagccgg 507
 Db 377 GGAGCCGACGACCGCGACGATCATCGGGGCTTCAGAAAGCGCTCGACGAGACCGG 436
 Qy 508 catgacggttcgatggccacacacacactcttcacccacccctcttcacagggcggcg 567
 Db 437 CCGTATCGTCCGATGGTGACCAACCACTCTTCAACCACCCGCTGTTCAAGGACGGCGG 496
 Qy 568 gttacacgccaacagcagcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 627
 Db 497 CTTACACGACGACCGCGCTTCCGTCGGCGCTACGCGATCCGCAAGGTGTCGCCAGAT 556
 Qy 628 cgaactcggtgagctgagtcggcgccaaaggtctacgtgcctggggcgccgagggcgcg 687
 Db 557 GGAGCTCGGCGCGGAGCTGGGCGGAGAGACGCTGCTCTTGGGGGCGCCGAGGGGCGC 616
 Qy 688 ggaactcgggtgcccacagagcgctgctgagccctgagccgcatgaagggagggccttga 747
 Db 617 CGAGTACGACTCGCGCAAGGACGTCAGCGCGCCCTTCGCGCTACCGCGAGGGCGCTCAA 676
 Qy 748 cctgctcggcgagtaagtcacactcgcgaggtcagacatcaggttcgacatcgagcccaa 807
 Db 677 CCGTCTCGCGAGCTACTCCGAGGACCGCGTTCGCGCTGCGCTTCGCCATTCGAGCCGAA 736
 Qy 808 gccgaacgagcgcgcgagcactcctgctccacacacacacacacacacacacacacacac 867
 Db 737 GCCAAGAGAGCCCGCGGCGACATCTCTCCGACCGCCGCGCCACCGCATCGCGTTCGT 796
 Qy 868 cgagcgcctgagcgcccgccgagctgacggtgtcaaaccccgaggtggggccacagagcagat 927
 Db 797 GCAGGAGCTGGAGCGTCCGAGCTCTTCGGCATCAACCCGAGAGACCGGGCACAGCAGAT 856
 Qy 928 ggcggcctgaaactcccgacgagcagcagcagcagcagcagcagcagcagcagcagcagc 987
 Db 857 GTCGAACCTCAACTTACCCAGGGCATCGCCAGCGCGCTGTCGACCAAGAAAGCTGTCCA 916
 Qy 988 catgacactcaacggccagtcgagcagcagcagcagcagcagcagcagcagcagcagcagc 1047
 Db 917 CATGACCTGACGGTTCAGCAGCGCCGCAAGTTCGACGAGGACCTGGTCTTCGGCCACGG 976
 Qy 1048 tgaactcgcgccgcttctgctggtgagcagcagcagcagcagcagcagcagcagcagcagc 1092
 Db 977 TGACTGTCTCAACGGGTCTTCGGTGTGACCTCTGAGAACGCGCCGCGAGCGGCCCC 1036
 Qy 1093 cggctggaggggtccgcccacttcgaactcaagcccgccgagccagggagcagcagcagcagc 1152
 Db 1037 GCGTACGAGCGGACCGCGCTCACTTCGACTACAAAGCCGCTCCGCTACCGAGGACTACGAGG 1096
 Qy 1153 cgtgtggcctcccgcgagcgaggtgcatgcgaactacatcctgagggagcgcgccgc 1212
 Db 1097 CGTCTGGAGTCGGGAGGAGGACCAATCCGATGTACCTGCTGCTCAAGAGGAGCGGCCAA 1156
 Qy 1213 cgccttcgctccgaccccgaggttcagagggcctgctgagcagcagcagcagcagcagcagc 1272

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: March 22, 2000, 23:30:34 ; Search time 27.38 Seconds
(without alignments)
333.924 Million cell updates/sec
Title: US-09-383-318-2
Perfect score: 2029
Sequence: 1 MNQPTPEDRFTGLTWGV.....GMAFERLDQLAMHLLGARG 386
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 188963 seqs, 23686106 residues
Total number of hits satisfying chosen parameters: 188963
Minimum DB seq length: 0
Maximum DB seq length: 1000000
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1900	93.6	388	1 R02215	Sequence encoding
2	1867.5	92.0	387	1 W63617	S. griseofuscus g1
3	1867.5	92.0	389	1 R02216	Sequence encoding
4	1867	92.0	388	1 P94687	DNA-deduced sequen
5	1863.5	91.8	387	1 W63617	S. griseofuscus g1
6	1859	91.6	387	1 R02217	Sequence encoding
7	1744.5	86.0	371	1 R03028	Xylose isomerase g
8	1715.5	84.5	345	1 R02218	Sequence encoding
9	1530.5	75.4	394	1 P90423	Streptomyces spp.
10	1361.5	67.1	394	1 R02213	Sequence encoding
11	1353.5	66.7	394	1 R03029	Xylose isomerase g
12	1351.5	66.6	394	1 R02212	Complete sequence
13	1348.5	66.5	394	1 R13280	Glucose isomerase
14	1347.5	66.4	394	1 R05282	Amino acid sequenc
15	1345.5	66.3	394	1 R13282	Glucose isomerase
16	1344.5	66.3	394	1 R13279	Glucose isomerase
17	1342.5	66.2	394	1 R13278	Glucose isomerase
18	1341.5	66.1	394	1 R13281	Glucose isomerase
19	1341.5	66.1	394	1 R13283	Glucose isomerase
20	1338.5	66.0	394	1 R13285	Glucose isomerase
21	1333.5	65.7	394	1 R13284	Glucose isomerase
22	1332.5	65.7	394	1 R13287	Glucose isomerase
23	1323.5	65.2	394	1 R13286	Glucose isomerase
24	1311.5	64.6	391	1 R13163	Glucose isomerase
25	1282	63.2	395	1 R02214	Sequence encoding
26	1280	63.1	395	1 R03030	Xylose isomerase g
27	1124	55.4	387	1 R22623	Thermus aquaticus
28	862	42.5	435	1 R44236	Glucose isomerase
29	320	15.8	444	1 W2508	Hyperthermostable
30	275.5	13.6	440	1 R22720	Xylose isomerase x
31	275	13.6	439	1 R03024	Xylose isomerase g
32	272.5	13.4	440	1 R05693	BglII fragment con
33	255.5	12.6	440	1 R03023	Xylose isomerase g
34	245.5	12.1	479	1 R99230	Barley xylose isom

35 106 5.2 3398 1 R44430
36 105 5.2 4572 1 W52845
37 103 5.1 573 1 R04715
38 99 4.9 572 1 R04716
39 99 4.9 573 1 R64765
40 96 4.7 800 1 P70420
41 95.5 4.7 329 1 W23389
42 95 4.7 921 1 W72033
43 94.5 4.7 573 1 R64766
44 94.5 4.7 878 1 W56116
45 94 4.6 1896 1 W72095

ALIGNMENTS

RESULT 1
R02215
ID R02215 standard; protein; 388 AA.
AC R02215;
DT 23-AUG-1990 (first entry)
DE Sequence encoding Streptomyces murinus DSM 40091 (Smu) glucose
DE isomerase (GI)
KW Glucose isomerase; Streptomyces murinus DSM 40091.
OS Streptomyces murinus.
PN Ep-351029-A.
PD 17-JAN-1990.
PF 17-JUL-1989; 201892.
PR 04-NOV-1988; Ep-402789.
PA (KONN) Gist-Brocades NV (Plan-).
PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;
DR WPI; 90-016368/03.
DR N-PSDB; Q93159.
PT New mutant glucose isomerase enzymes -
PT obtd. by site-directed mutagenesis of gene from Actinoplanes
PT missouriensis, used for prodn. of high fructose corn syrups
PS Disclosure; 56pp; English.
CC The invention is a novel mutant glucose isomerase (GI) enzyme with
CC improved properties thru AA substitn. The GI is pref. derived from
CC Actinoplanes missouriensis. The substitn. is Lys for Arg.
CC or vice versa. Figure 21 compares AA sequences of GI from different
CC sources.
SQ Sequence 388 AA;

Query Match 93.6%; Score 1900; DB 1; Length 388;
Best Local Similarity 92.8%; Pred. No. 3.4e-171;
Matches 360; Conservative 14; Mismatches 12; Indels 2; Gaps 2;
QY 1 MNQPTPEDRFTGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
Db 1 MSFQPTPEDRFTGLTWGVQGRDPFGDTRPALDPVETVQRLAELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVKKRFRQALDAGTGMTPMATTNLFTHPVFKAGFTANDRAVRVYALRKTI 119
Db 61 FGSSDTERESHKRRFQALDAGTGMTPMATTNLFTHPVFDGFTANDRVRYALRKTI 120
QY 120 RNIDLAELGAKYVYAWGREGAEGAAKDVRAALDRMKEAFDGLLGEYVTSQGYDIRFAI 179
Db 121 GNIDLAELGAKYVYAWGREGAEGAAKDVRAALDRMKEAFDGLLGEYVTAQGYDIRFAI 180
QY 180 EPKPNPRGDIILLPTIGHALAFIERLERPELVGNPVEVGEHQMAGLNFPFHGIAQALWAGK 239
Db 181 EPKPNPRGDIILLPTVGHALAFIERLERPELVGNPVEVGEHQMAGLNFPFHGIAQALWAGK 240
QY 240 LPHIDLNGSGIKYDODLRFAGDLRAAFWLYDVLLESAGWEGPRHDFKPPRTEDIDGVW 299
Db 241 LPHIDLNGSGIKYDODLRFAGDLRAAFWLYDVLLESAGWEGPRHDFKPPRTEDIDGVW 300
QY 300 ASAAGCMRNLYLTKERAAAFRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDF 359
Db 301 ASAAGCMRNLYLTKERAAAFRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDF 360

eryA region polype
A. mediterranei r1
Amino acid sequenc
Amino acid sequenc
M. leprae 65 kDa p
Sequence encoded b
Shaker-like potass
HSV-2 strain SB5 C
M. tuberculosis 65
Microbispora therm
HSV-2 strain SB5 C

PI Drummond RJ, Bloch W, Matthews BW, Toy PL, Nicholson HH;
 DR WPI; 89-068875/09.
 N-PSDB; N91083.
 PT Increasing stability of proteins by specific aminoacid replacement -
 PT where aminoacid introduced decreases configuration entropy of
 PT unfolding of protein
 PS Disclosure: ; 134pp; English.
 CC New in the patent are mutans of the native protein given in p94687.
 CC These mutans are used to convert glucose to fructose, and xylose to
 CC xylulose, eg in mfr. of sweeteners. Compared with native xylose the
 CC mutans have different chemical and thermal stability, kinetic constants,
 CC specificity and/or lower optimum pH. Amino acid substitution sites are
 CC selected by first determining, from the crystallographic structure, the
 CC phi and psi angles, then screening for these for values within the specified
 CC ranges eg a site which has a phi backbone configuration angle -40 to
 CC -90 degrees when psi configuration angle is 0 to -60 degrees, or -40 to
 CC 95 degrees when psi - 120 to 180 degrees, and is able to accommodate
 CC AA without distortion of the 3-D protein structure, so that the
 CC substitution decreases the configurational entropy of unfolding CEU; or
 CC replacing a Gly residue having a negative phi angle with Ala, also
 CC decreasing the CEU.
 SQ Sequence 388 AA;

Query Match 92.0%; Score 1867; DB 1; Length 388;
 Best Local Similarity 91.8%; Pred. No. 4.4e-168;
 Matches 356; Conservative 14; Mismatches 16; Indels 2; Gaps 2;

QY 1 MNYQTPEDRFTFGLTWGQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDLLIP 59
 DB 1 MNYQTPEDRFTFGLTWGQGRDPFGDTRPALDPVESVRLAELGAYGVTFHDDLLIP 60
 QY 60 FGASDERAHVKRFRQALDGMTVPMTNLFTHPVFKAGFTANDRVRRYALRKT 119
 DB 61 FGSSDEREHHVKRFRQALDGMTVPMTNLFTHPVFKAGFTANDRVRRYALRKT 120
 QY 120 RNIDLAELGAYGVYVANGREGAESGAKDVRAALDRMKEAFDGLGEYTSOGYDIRFAI 179
 DB 121 RNIDLAELGAYGVYVANGREGAESGAKDVRAALDRMKEAFDGLGEYTSOGYDIRFAI 180
 QY 180 EPKPNPRGDILLPTIGHALAFIERLERPELVGNPEVGEQMGALNFPHGIAQALWACK 239
 DB 181 EPKPNPRGDILLPTIGHALAFIERLERPELVGNPEVGEQMGALNFPHGIAQALWACK 240
 QY 240 LPHIDLNGSGIKYDQDLRFAGADLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGW 299
 DB 241 LPHIDLNGSGIKYDQDLRFAGADLRAAFWLVLDLESAGYSGPRHDFKPPRTEDIDGW 300
 QY 300 ASAAGCMRNYLLKERRAAFRADPEVQEARLRAALDOLAEPTAADGLQALLADRTAYEDF 359
 DB 301 ASAAGCMRNYLLKERRAAFRADPEVQEARLRAALDOLAEPTAADGLHPLDDRSFAFEF 360
 QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
 DB 361 DVDAAARGMAFERLDQALMDHLLGARG 388

RESULT 5
 ID W63618 standard; Protein; 387 AA.
 AC W63618;
 DT 21-SEP-1998 (first entry)
 DE S. griseofuscus glucose isomerase mutant Y252F.
 KW Glucose isomerase; heat resistance; thermostable; mutant;
 OS Isomerised sugar.
 OS Streptomyces griseofuscus.
 OS Synthetic.
 FT Key Location/Qualifiers
 FT Misc_difference 253 /label= Y252F
 FT /note= "mature wild-type Tyr is replaced with Phe"
 PN J10084956-A.
 PD 07-APR-1998.

PF 12-SEP-1996; 262370.
 PR 12-SEP-1996; JP-262370.
 PA (GODO) GODO SHUSEI KK.
 DR WPI; 98-264846/24.
 PT Improving heat resistance of glucose isomerase - comprises replacing
 PT part of amino acid in specific position of amino acid sequence
 PS Claim 1; Page -; 7pp; Japanese.
 CC This represents a mutant of Streptomyces glucose isomerase that has
 CC improved heat resistance. The mutant was created by replacing an
 CC amino acid in a specific position of the mature wild-type sequence.
 CC The glucose isomerase is useful for the preparation of isomerised
 CC sugar.
 CC Note: this sequence is not provided in the specification; it has been
 CC created by modifying the glucose isomerase sequence provided in Fig 1.
 SQ Sequence 387 AA;

Query Match 91.8%; Score 1863.5; DB 1; Length 387;
 Best Local Similarity 91.5%; Pred. No. 9.4e-168;
 Matches 355; Conservative 17; Mismatches 13; Indels 3; Gaps 3;

QY 1 MNYQTPEDRFTFGLTWGQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDLLIP 59
 DB 1 MSFQTPEDKFTFGLTWGQGRDPFGDTRPALDPVETVQRLAELGAYGVTFHDDLLIP 60
 QY 60 FGASDERAHVKRFRQALDGMTVPMTNLFTHPVFKAGFTANDRVRRYALRKT 119
 DB 61 FGSSDERESHKFRFRQALDGMTVPMTNLFTHPVFK-DRFTANDRVRAYVRKTI 119
 QY 120 RNIDLAELGAYGVYVANGREGAESGAKDVRAALDRMKEAFDGLGEYTSOGYDIRFAI 179
 DB 120 RNIDLAELGAYGVYVANGREGAESGAKDVRAALDRMKEAFDGLGEYTAGYDIRFAI 179
 QY 180 EPKPNPRGDILLPTIGHALAFIERLERPELVGNPEVGEQMGALNFPHGIAQALWACK 239
 DB 180 EPKPNPRGDILLPTIGHALAFIERLERPELVGNPEVGEQMGALNFPHGIAQALWACK 239
 QY 240 LPHIDLNGSGIKYDQDLRFAGADLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGW 299
 DB 240 LPHIDLNGSGIKYDQDLRFAGADLRAAFWLVLDLESAGYEGPRHDFKPPRTEDIDGW 299
 QY 300 ASAAGCMRNYLLKERRAAFRADPEVQEARLRAALDOLAEPTAADGLQALLADRTAYEDF 359
 DB 300 ASAAGCMRNYLLKERRAAFRADPEVQEARLRAALDOLAEPTAADGLQALLADRTAFEDF 359
 QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
 DB 360 DVDAAARGMAFERLDQALMDHLLGARG 387

RESULT 6
 ID R02217 standard; protein; 387 AA.
 AC R02217;
 DT 23-AUG-1990 (first entry)
 DE Sequence encoding Streptomyces violaceoruber LMG 7183 (Svr) glucose
 DE isomerase (GI)
 KW Glucose isomerase; Streptomyces violaceoruber.
 OS Streptomyces violaceoruber.
 PN EP-351029-A.
 PD 17-JAN-1990.
 PF 17-JUL-1989; 201892.
 PR 04-NOV-1988; EP-402789.
 PA (KONN) Gist-Brocades NV (Plan-).
 PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;
 DR WPI; 90-016368/03.
 PT New mutant glucose isomerase enzymes -
 PT obtd. by site-directed mutagenesis of gene from Actinoplanes
 PT missouriensis, used for prodn. of high fructose corn syrups
 PS Disclosure; ; 56pp; English.
 CC The invention is a novel mutant glucose isomerase (GI) enzyme with
 CC improved properties thru AA substitn. The GI is pref. derived from
 CC Actinoplanes missouriensis. The substitn. is Lys for Arg.

```
CC or vice versa. Figure 21 compares AA sequences of GI from different
CC sources.
SQ Sequence 387 AA;

Query Match 91.6%; Score 1859; DB 1; Length 387;
Best Local Similarity 91.2%; Pred. No. 2.5e-167;
Matches 353; Conservative 15; Mismatches 17; Indels 2; Gaps 2;

QY 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
DQ 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVKKRFRQALDQATGTVPMATTNLFTHPVFKAGFTANDRAVRRYALRKT 119
DQ 61 FGSDTERESHKFRQALDQATGTVPMATTNLFTHPVFKAGFTANDRAVRRYALRKT 120
QY 120 RNIDLAVELGAKYVANGREGSAGAAKDVRAALDRMKEAFDGLLGEYVTSQGYDIRFAI 179
DQ 121 RNIDLAVELGASVYVANGREGSAGAAKDVRAALDRMKEAFDGLLGEYVTSQGYDLKFAI 180
QY 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
DQ 181 EPKNEPRGDILLPTVGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 240
QY 240 LFHIDLNGSQGIKYDQDLRFAGDGLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 299
DQ 241 LFHIDLNGSQGIKYDQDLRFAGDGLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 300
QY 300 ASAAGCHRNLYILKERAARFADPEVQALRAARLDQAEPTAAGDGLQALLADRTAYEDF 359
DQ 301 ASAAGCHRNLYILKERAARFADPEVQALRAARLDQAEPTAAGDGLQALLADRTAYEDF 360
QY 360 DVD-AAAGMFAFERLDQLAMDHLLGARG 385
DQ 361 DVDAAGMFAFERLDQLAMDHLLGARG 387

RESULT 7
ID R03028 standard; protein; 371 AA.
AC R03028;
DE 11-JUL-1990 (first entry)
KW Xylose isomerase gene of Streptomyces violaceoniger.
OS Xylose isomerase; Arthobacter.
PN W09000198-A.
PF 04-JUL-1989; G00748.
PR 04-JUL-1988; GB-015902.
PI (BLOW) Blow D M.
PI Blow DM, Hartley BS, Henrick K;
DR WPI; 90-037131/05.
PT Xylose isomerase mutants -
PT having amino acid replacements to improve activity under acid pH
PT conditions and/or increase stability and/or affinity for metal ions.
PS Disclosure; Fig 7; 35pp; English.
CC Mutants of the xylose isomerase (XI) gene esp. Arthobacter strain B3728,
CC have been shown to have increased thermal stability, and/or improved
CC activity at low pH, and/or affinity for metal ions esp. Ca.
CC Useful in industrial conversion of glucose to sweeter sugars.
SQ Sequence 371 AA;

Query Match 85.08; Score 1744.5; DB 1; Length 371;
Best Local Similarity 87.2%; Pred. No. 1.5e-156;
Matches 340; Conservative 18; Mismatches 9; Indels 23; Gaps 6;

QY 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
DQ 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVKKRFRQALDQATGTVPMATTNLFTHPVFKAGFTANDRAVRRYALRKT 119

CC or vice versa. Figure 21 compares AA sequences of GI from different
CC sources.
SQ Sequence 387 AA;

Query Match 84.5%; Score 1715.5; DB 1; Length 345;
Best Local Similarity 93.6%; Pred. No. 7.1e-154;
Matches 323; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIP 59
DQ 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVKKRFRQALDQATGTVPMATTNLFTHPVFKAGFTANDRAVRRYALRKT 119
DQ 61 FGAEDEAREAHVKKRFRQALDQATGTVPMATTNLFTHPVFKAGFTANDRAVRRYALRKT 120
QY 120 RNIDLAVELGAKYVANGREGSAGAAKDVRAALDRMKEAFDGLLGEYVTSQGYDIRFAI 179
DQ 121 RNIDLAVELGARTYVANGREGSAGAAKDVRAALDRMKEAFDGLLGEYVTSQGYDIRFAI 180
QY 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
DQ 181 EPKNEPRGDILLPTVGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 240
QY 240 LFHIDLNGSQGIKYDQDLRFAGDGLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 299
```

```

|||||
Db 241 LFHIDNGSGIYKQDRLFGAGDLRAAFWLVLDLSSGYDGRPHDFDKPRTEDLDGVW 300
QY 300 ASAAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPRTAAD 344
|||||
Db 301 ASAAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPRTAAD 345

RESULT 9
P90423
ID P90423 standard; protein; 394 AA.
AC P90423;
DT 1-NOV-1989 (first entry)
DE Streptomyces spp. glucose isomerase
KW Glucose isomerase; Streptomyces; DNA.
OS Streptomyces
PN J01137979-A.
PD 24-MAY-1989.
PF 24-NOV-1987; 295739.
PR 24-NOV-1987; JP-295739.
PA (NORO) Norinsho.
DR WPI; 89-198224/27.
DR N-PSDB; N90362.
PT Novel glucose isomerase gene
PT - exists in chromosome of Streptomyces and is
PT contained in recombinant etc.
PS Claim 1; fig 1; 6pp; Japanese.
CC Glucose isomerase from Streptomyces spp. (see N90362). Used to
CC study this protein's structure by recombinant techniques, and for
CC prodn. of an enzyme with improved properties.
SQ Sequence 394 AA;

Query Match 75.4%; Score 1530.5; DB 1; Length 394;
Best Local Similarity 79.0%; Pred. No. 2.3e-136;
Matches 305; Conservative 21; Mismatches 25; Indels 35; Gaps 8;

QY 1 MNYQPTPEDRFTGLTWGVQGRDPFGDATRPALDPVD-VQRLAELGAYGVTHDDDLIP 59
|||||
Db 1 MSQPTPEDKFTGLTWGVQGRDPFGDATRPGLDPVETVQRLAELGAYGVTHDDDLNP 60
|||||
QY 60 FGASDTEREAHVKKRQALDAGTMTVPMTNLTFTHPVKAGAFANDRAVRYALRKTII 119
|||||
Db 61 FGSDTEREASHIRKQALDAGTMTVPMTNLTFTHPVKAF-DRTANDRDVAYAVRKTII 119
|||||
QY 120 RNIDLAELGAKYVAVWGREGAESAAGKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 179
|||||
Db 120 RNIDLAELGAKYVAVWGREGAESAAGKDVRAALDRMKEAFDLLGEYVTAQGYDLRFAI 179
|||||
QY 180 EKPNEPRGDIILLPTGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239
|||||
Db 180 EKPNEPRGDIILLPTGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239
|||||
QY 240 LFHIDNGSGIYKQDRL---RGAGDLRAAFWLVLDLSSAGHEGRPHDFDKPRTEDID 296
|||||
Db 240 LFHIDNGSGIYKQDGRRRPAGGV-----LVVLDLSSAGYEGRPHDFDKPRTEDFD 295
|||||
QY 297 GWASAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPRTAAD 343
|||||
Db 296 GWASAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPRTAAD 343
|||||
QY 344 DGLQALLADRTAYEDFDVD-AAARGM 368
|||||
Db 344 DGLQALLADRTAFEDVDEAAARGM 369

RESULT 10
R02213
ID R02213 standard; protein; 394 AA.
AC R02213;
DT 23-AUG-1990 (first entry)
DE Sequence encoding Ampullariella species ATCC 31351 glucose
DE isomerase (GI)

```

```

KW Glucose isomerase; Ampullariella species ATCC 31351.
OS Ampullariella.
PN EP-351029-A.
PD 17-JAN-1990.
PF 04-JUL-1989; 201892.
PR 04-NOV-1988; EP-402789.
PA (KONN) Gist-Brocades NV (Plan-).
PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;
DR WPI; 90-016368/03.
PT New mutant glucose isomerase enzymes -
PT obtd. by site-directed mutagenesis of gene from Actinoplanes
PT missouriensis, used for prodn. of high fructose corn syrups
PS Disclosure; 56pp; English.
CC The invention is a novel mutant glucose isomerase (GI) enzyme with
CC improved properties thru AA substitn. The GI is pref. derived from
CC Actinoplanes missouriensis. The substitn. is Lys for Arg.
CC or vice versa. Figure 21 compares AA sequences of GI from different
CC sources.
SQ Sequence 394 AA;

Query Match 67.1%; Score 1361.5; DB 1; Length 394;
Best Local Similarity 66.5%; Pred. No. 1.9e-120;
Matches 262; Conservative 45; Mismatches 78; Indels 9; Gaps 4;

QY 1 MNYQPTPEDRFTGLTWGVQGRDPFGDATRPALDPVD-VQRLAELGAYGVTHDDDLIP 59
|||||
Db 1 MSQATPDCKFSFGLTWGVQGRDAFGDATRPVLDPEAVHKLAEIGAYGVTHDDDLVP 60
|||||
QY 60 FGASDTEREAHVKKRQALDAGTMTVPMTNLTFTHPVKAGAFANDRAVRYALRKTII 119
|||||
Db 61 FGADAATRDGIAGVFSKALDETGLIVPMVTNLTFTHPVKDGGFTSNDRSVRYAIRKVL 120
|||||
QY 120 RNIDLAELGAKYVAVWGREGAESAAGKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 179
|||||
Db 121 RQMDLGAELGAKTLVWGREGAEDSANDVGAALDRYREALNLLAQYSEDQGYGLRFAI 180
|||||
QY 180 EKPNEPRGDIILLPTGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239
|||||
Db 181 EKPNEPRGDIILLPTAGHAIAFVQELERPELFGINPEHQMSNLTQTGIAQALWHKK 240
|||||
QY 240 LFHIDNGSGIYKQDRLFGAGDLRAAFWLVLDLSSA-----GWEGPRHDFDKPRTED 294
|||||
Db 241 LFHIDNGSGIYKQDGLVFGHGLDLNLSFVLDLLENGPDGGPAYDGRPHDFDKPRTED 300
|||||
QY 295 IDGVASAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPRTAADG--LQALLAD 352
|||||
Db 301 FDGVWESAKDNIRMYLLILKERAKAFRAADPEVQEAALRAARLDQLAEPRTAAD 360
|||||
QY 353 RTAYEDFDVDA-AAARGMAFERLDQLAHDHLLGAR 385
|||||
Db 361 RSFADYDADAVGAKGYGVFKLNQLAIDLHLLGAR 394

RESULT 11
R03029
ID R03029 standard; protein; 394 AA.
AC R03029;
DT 11-JUL-1990 (first entry)
DE Xylose isomerase gene of Ampullariella sp.
KW Xylose isomerase; Arthobacter.
OS Ampullariella sp. strain 3876.
PN W09000196-A.
PD 11-JAN-1990.
PF 04-JUL-1989; G00748.
PR 04-JUL-1988; GB-015902.
PI Blow DM, Hartley BS, Henrick K;
DR WPI; 90-037131/05.
PT Xylose isomerase mutants -
PT having amino acid replacements to improve activity under acid pH
PT conditions and/or increase stability and/or affinity for metal ions.
PS Disclosure; Fig 7; 35pp; English.

```

CC Mutants of the xylose isomerase (XI) gene esp. *Arthobacter* strain B3728,
 CC have been shown to have increased thermal stability, and/or improved
 CC activity at low pH, and/or affinity for metal ions esp. Ca.
 CC Useful in industrial conversion of glucose to sweeter sugars.
 SQ Sequence 394 AA;

```

Query Match      66.7%; Score 1353.5; DB 1; Length 394;
Best Local Similarity 66.0%; Pred. No. 1.1e-119;
Matches 260; Conservative 46; Mismatches 79; Indels 9; Gaps 4;

Qy 1 MNYQTPEDRFTFGWTVGMQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
Db 1 MSLQATPDCKFSFGLWTVGMQARDAFGDAFTRVLPDPIEAVHKLAEIGAYGVTFHDDDLVP 60
Qy 60 FGASDTEREAHVKRFQALDAGMTVPMTTNLFTHPVFKAGAFANDRAVRYALRKTI 119
Db 61 FGDAATRDGIVAGFSKALDETGLIVPMVTTLNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120
Qy 120 RNIDLAVALGAKVYVAMGREGAESAAGKADYRAALDRKKEAFDILGGEYVTSQGYDIRFAI 179
Db 121 RQMDLGAELGAKTLVLMGREGAEYDSAKDVGAALDRYREALNLLAQYSEDQGYGLPFAI 180
Qy 180 EPKNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPHPHGIAQALWAGK 239
Db 181 EPKNEPRGDIILLPTAGHAI AFVQELERPELFGINPETGHEQMSNLNFTQGIQAALWHKK 240
Qy 240 LFHIDLNGSGIKYDODLRFAGGDLRAAFWLVLDLLESA-----GWEGPRHFDKPPRTED 294
Db 241 LFHIDLNGHGPCKDODLVFGHGDLNLFSLVDLLENGPDGPAVDGPRHFDYKPSRTED 300
Qy 295 IDGVWASAAGCMRNLYILKERAARADPEVQEAALRAARDQLAEPTADG--LQALLAD 352
Db 301 FDGVWESAKDNIRMYLLKRAKAFRADPEVQEAALAEKSVDELRTPTLNPGETYADLLAD 360
Qy 353 RTAYEDFDVDA-AARGMAFERLDQLAMDHLLGAR 385
Db 361 RSAFEDYDADAVGAKGFGVKLNQLAIDHLLGAR 394

```

```

RESULT 12
R02212
ID R02212 standard; protein; 394 AA.
AC R02212;
DT 23-AUG-1990 (first entry)
DE Complete sequence of wildtype (WT) Actinoplanes missouriensis
DE Glucose isomerase (GI)
KW Site-directed mutagenesis; glucose isomerase; Actinoplanes missouriensis;
KW high fructose corn syrup.
OS Actinoplanes missouriensis.
PN EP-351029-A.
PD 17-JAN-1990.
PF 17-JUL-1989; 201892.
PR 04-NOV-1988; EP-402789.
PA (KONN) Gist-Brocades NV (Plan-).
PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;
DR N-PSDB; Q93158.
PT New mutant glucose isomerase enzymes -
PT obtd. by site-directed mutagenesis of gene from Actinoplanes
PT missouriensis, used for prodn. of high fructose corn syrups
PS Disclosure; 56pp; English.
CC The invention is a novel mutant glucose isomerase (GI) enzyme with
CC improved properties thru AA substitn. The GI is pref. derived from
CC Actinoplanes missouriensis. The substitn. is Lys for Arg.
CC or vice versa. Specific Lys or Arg residues are identified and site-
CC directed mutagenesis of the DNA sequence encoding the GI is performed.
CC The novel mutant gi shows 65 per cent or more sequence homology with the
CC AA sequence of WT A. missouriensis GI. Compared to wt enzyme it shows
CC higher conversion performance and improved thermostability and pH
CC stability. 394 AA;
SQ Sequence

```

```

Query Match      66.6%; Score 1351.5; DB 1; Length 394;
Best Local Similarity 65.5%; Pred. No. 1.7e-119;
Matches 258; Conservative 51; Mismatches 76; Indels 9; Gaps 4;

Qy 1 MNYQTPEDRFTFGWTVGMQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
Db 1 MSVQATREDKFSFGLWTVGMQARDAFGDAFTRVLPDPIEAVHKLAEIGAYGVTFHDDDLVP 60
Qy 60 FGASDTEREAHVKRFQALDAGMTVPMTTNLFTHPVFKAGAFANDRAVRYALRKTI 119
Db 61 FGSAQTRDGIAGFKALDETGLIVPMVTTLNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120
Qy 120 RNIDLAVALGAKVYVAMGREGAESAAGKADYRAALDRKKEAFDILGGEYVTSQGYDIRFAI 179
Db 121 RQMDLGAELGAKTLVLMGREGAEYDSAKDVSAALDRYREALNLLAQYSEDQGYGLPFAI 180
Qy 180 EPKNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPHPHGIAQALWAGK 239
Db 181 EPKNEPRGDIILLPTAGHAI AFVQELERPELFGINPETGHEQMSNLNFTQGIQAALWHKK 240
Qy 240 LFHIDLNGSGIKYDODLRFAGGDLRAAFWLVLDLLESA-----GWEGPRHFDKPPRTED 294
Db 241 LFHIDLNGHGPCKDODLVFGHGDLNLFSLVDLLENGPDGPAVDGPRHFDYKPSRTED 300
Qy 295 IDGVWASAAGCMRNLYILKERAARADPEVQEAALRAARDQLAEPT--AADGLQALLAD 352
Db 301 YDGVWESAKANIRMYLLKRAKAFRADPEVQEAALAEKSVDELRTPTLNPGEYAEILLAD 360
Qy 353 RTAYEDFDVDA-AARGMAFERLDQLAMDHLLGAR 385
Db 361 RSAFEDYDADAVGAKGFGVKLNQLAIEHLLGAR 394

```

```

RESULT 13
R13280
ID R13280 standard; Protein; 394 AA.
AC R13280;
DT 16-OCT-1991 (first entry)
DE Glucose isomerase mutant A25K.
KW GI; high fructose corn syrup; sweetener; sugar.
OS Actinoplanes missouriensis.
FH Key Location/Qualifiers
FT domain 10..14
FT /label= beta_strand
FT domain 49..52
FT /label= beta_strand
FT domain 87..91
FT /label= beta_strand
FT domain 132..137
FT /label= beta_strand
FT domain 177..181
FT /label= beta_strand
FT domain 211..218
FT /label= beta_strand
FT domain 241..246
FT /label= beta_strand
FT domain 289..292
FT /label= beta_strand
FT domain 35..47
FT /label= helix
FT domain 64..83
FT /label= helix
FT domain 108..130
FT /label= helix
FT domain 150..173
FT /label= helix
FT domain 195..204
FT /label= helix
FT domain 127..239
FT /label= helix
FT domain 264..276
FT /label= helix

```

```

FT domain 300..327
FT FT /label= helix
FT region 25
FT FT /label= mutation
FT FT /note=" Ala -> Lys "
PN EP-440273-A.
PD 07-AUG-1991.
PF 02-JAN-1991; 200003.
PR 04-JAN-1990; EP-200029.
PR 02-JAN-1991; EP-200003.
PA (KONN ) GIST-BROCADES NV.
PA (PLAN-) PLANT GENETIC SXST NV.
PI Lambeir AVR, Quax WJ, Lasters I, Van der Laan JM;
DR WPI: 91-231993/32.
PT Glucose isomerase mutants with altered substrate specificity -
PT and methods for selecting aminoacid(s) for substitution, used
PT e.g. in prodn. of high fructose corn syrup.
PS Claim 13; Page 13; 18pp; English.
CC The analogue was prepd. by site directed mutagenesis of the GI
CC gene. The substn. of Lys for Ala at posn 25 introduces a positive
CC charge at a distance of 6-8A of the O1, O2 and O3 positions of the
CC substrate in the enzyme-orbitol-cobalt complex. This causes
CC disruption of the water structure in the interface and displaces
CC 26pe, which shapes the hydrophobic pocket accommodating the C1
CC aliphatic hydrogens of the substrate. This gives the enzyme a
CC better Km(xylose)/Km(glucose) ratio compared with the wild type.
CC The GI is used in industrial processes, e.g. for the prodn. of
CC high fructose corn syrup.
CC See also R13278-R13287.
SQ Sequence 394 AA;

Query Match 66.5%; Score 1348.5; DB 1; Length 394;
Best Local Similarity 65.5%; Pred. No. 3.2e-119;
Matches 258; Conservative 50; Mismatches 77; Indels 9; Gaps 4;

QY 1 MNYQPTPDRFTFGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
DB 1 MSVQATREDKFSFGLTWGVQGRDPFGDTRPALDPVEAVHKLAEIGAYGVTFHDDDLIP 60
QY 60 FGASDERAHVHVRFRQALDGMTVPMTNLTFTHPVKAGAFANDRAVRYALRKTI 119
DB 61 FGSDAQTRDGIAGFKKALDEGLIVPMVTNLTFTHPVFDGFTSNDRSVRYAIRKVL 120
QY 120 RNIDLAVELGAKVYVWAGREGAESGAADVRAALDRKAEFDLLGEYVTSOGYDIRFAI 179
DB 121 RQMDLGAELGAKTLVWGREGAEYDSAKDVSAALDRYREALNLLAQYSEDGRYGLRFAI 180
QY 180 EPKPNPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239
DB 181 EPKPNPRGDIILLPTAGHAIQVQLERPELEFGINPEIGHEQMSNLFNFTQGIQAALWHRK 240
QY 240 LFHIDLNGSGIKYDQDLRFAGAGDLRAAFWLVDDLESA-----GWEGPRHFDKPPRTED 294
DB 241 LFHIDLNGHGGPKFDQDLVFGHGDLNLSLVDDLENGPDGAPAYDGPGRHFDKPPRTED 300
QY 295 IDGVWASAGCMRNLYLILKERAAFRADPEVQEAALRAALDQLAEPT--AADGLQALLAD 352
DB 301 YDGVWESAKANIRMYLLKLERAKAFRADPEVQEAALAAKVAELKPTLNPGEYAEALLAD 360
QY 353 RTAYEDFDVDA-AARGMAFERLDQLAMDHLLGAR 385
DB 361 RSAFEDYDADAVGAKGFGFKLNQLNLAIEHLLGAR 394

RESULT 14
R05282
ID R05282 standard; protein; 394 AA.
AC R05282;
DT 20-AUG-1990 (first entry)
DE Amino acid sequence of Actinoplanes missouriensis (DSN 43046)
DE D-glucose isomerase (EcoAmi(DSM) GI)
KW Actinoplanes missouriensis (DSM 43046); D-glucose isomerase;

KW D-xylose isomerase; D-xylose.ketol-isomerase.
OS Actinoplanes missouriensis.
PN EP355039-A.
PD 21-FEB-1990.
PF 17-JUL-1989; 201893.
PR 17-JUL-1989; EP-201893, EP-201539.
PA (KONN) Gist-Brocades NV (PLAN-).
PI Mrabet N, Lasters I, Stanssens P, Mathysens G, Wodak S, Quax WJ;
DR WPI: 90-052789/08.
DR N-PSDB; Q01862.
PT Biologically active modified protein prodn. -
PT comprises substituting arganine residue for lysine in starting protein
PS Disclosure; Fig 15; 64pp; English.
CC It is D-glucose isomerase from Actinoplanes missouriensis
CC (DSM 43046). Its DNA is genetically engineered to code for a mutein
CC that has increased thermostability and increased resistance
CC towards chemical modification. Its DNA is used in an example of
CC the method, which is claimed, where one or more Lys are substd. by Arg or
CC vice versa, at a site which can accommodate such a substitution without
CC altering the biological activity.
SQ Sequence 394 AA;

Query Match 66.4%; Score 1347.5; DB 1; Length 394;
Best Local Similarity 65.2%; Pred. No. 4e-119;
Matches 257; Conservative 52; Mismatches 76; Indels 9; Gaps 4;

QY 1 MNYQPTPDRFTFGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
DB 1 VSVQATREDKFSFGLTWGVQGRDPFGDTRPALDPVEAVHKLAEIGAYGVTFHDDDLIP 60
QY 60 FGASDERAHVHVRFRQALDGMTVPMTNLTFTHPVKAGAFANDRAVRYALRKTI 119
DB 61 FGSDAQTRDGIAGFKKALDEGLIVPMVTNLTFTHPVFDGFTSNDRSVRYAIRKVL 120
QY 120 RNIDLAVELGAKVYVWAGREGAESGAADVRAALDRKAEFDLLGEYVTSOGYDIRFAI 179
DB 121 RQMDLGAELGAKTLVWGREGAEYDSAKDVSAALDRYREALNLLAQYSEDGRYGLRFAI 180
QY 180 EPKPNPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239
DB 181 EPKPNPRGDIILLPTAGHAIQVQLERPELEFGINPEIGHEQMSNLFNFTQGIQAALWHRK 240
QY 240 LFHIDLNGSGIKYDQDLRFAGAGDLRAAFWLVDDLESA-----GWEGPRHFDKPPRTED 294
DB 241 LFHIDLNGHGGPKFDQDLVFGHGDLNLSLVDDLENGPDGAPAYDGPGRHFDKPPRTED 300
QY 295 IDGVWASAGCMRNLYLILKERAAFRADPEVQEAALRAALDQLAEPT--AADGLQALLAD 352
DB 301 YDGVWESAKANIRMYLLKLERAKAFRADPEVQEAALAAKVAELKPTLNPGEYAEALLAD 360
QY 353 RTAYEDFDVDA-AARGMAFERLDQLAMDHLLGAR 385
DB 361 RSAFEDYDADAVGAKGFGFKLNQLNLAIEHLLGAR 394

RESULT 15
R13282
ID R13282 standard; Protein; 394 AA.
AC R13282;
DT 16-OCT-1991 (first entry)
DE Glucose isomerase mutant T90S.
KW GI; high fructose corn syrup; sweetener; sugar.
OS Actinoplanes missouriensis.
FH Key Location/Qualifiers
FT domain 10..14
FT FT /label= beta_strand
FT domain 49..52
FT FT /label= beta_strand
FT domain 87..91
FT FT /label= beta_strand
FT domain 132..137
FT FT /label= beta_strand

```

FT domain 177..181
 FT /label= beta_strand
 FT domain 211..218
 FT /label= beta_strand
 FT domain 241..246
 FT /label= beta_strand
 FT domain 289..292
 FT /label= beta_strand
 FT domain 35..47
 FT /label= helix
 FT domain 64..83
 FT /label= helix
 FT domain 108..130
 FT /label= helix
 FT domain 150..173
 FT /label= helix
 FT domain 195..204
 FT /label= helix
 FT domain 127..239
 FT /label= helix
 FT domain 264..276
 FT /label= helix
 FT domain 300..327
 FT /label= helix
 FT region 186
 FT /label= mutation
 FT /note=" Glu -> Gln "
 PN EP-440273-A.
 PD 07-AUG-1991.
 PF 02-JAN-1991; 200003.
 PR 04-JAN-1990; EP-200029.
 PR 02-JAN-1991; EP-200003.
 PA (KONN) GIST-BROCADES NV.
 PA (PLAN-) PLANT GENETIC SYST NV.
 PI Lambelir AVR, Quax WJ, Lasters I, Van der Leen JM;
 DR WPI; 91-231993/32.
 PT Glucose isomerase mutants with altered substrate specificity -
 PT and methods for selecting aminoacid(s) for substitution, used
 PT e.g. in prodn. of high fructose corn syrup.
 PS Claim 13; Page 13; 18pp; English.
 CC The analogue was prepd. by site directed mutagenesis of the GI
 CC gene. The substn. of Glu for Glu at posn 186 gives the enzyme
 CC a better Km(xylose)/Km(glucose) ratio than the wild type.
 CC The GI is used in industrial processes, e.g. for the prodn. of
 CC high fructose corn syrup.
 CC See also R13278-R13287.
 SQ Sequence 394 AA;

Query Match 66.3%; Score 1345.5; DB 1; Length 394;
 Best Local Similarity 65.2%; Pred. No. 6.2e-119;
 Matches 257; Conservative 51; Mismatches 77; Indels 9; Gaps 4;
 QY 1 MNYQTPEDRFTFTGLWTVGMQRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
 DB 1 MSVQATREDKFSFGLTWGQARDATRTALDPVEAVHKLAEIGAYGTTFHDDDLVP 60
 QY 60 FGASDTERAHVRRQALDAGTGMVPMATNTLFTHPVKAGFTANDRAVRYALRKTI 119
 DB 61 FGSDAQRDGIAGFKKALDETGLIVPMVTNTLFTHPVKGFGFTSDNSRVRYAIRKVL 120
 QY 120 RNIDLAVELGAKVYVAMGREGAEGSAKDVRAALDRKAEFDLLGEVYTSQGYDIRFAI 179
 DB 121 RQMDLGAELGAKTLVLMGREGAEGYDSAKDVSAAALDRYREALNLLAQYSEDGRGYGLRFAI 180
 QY 180 EPKPNPRGDIILLPTIGHALAFIERLERPELYGVNPEVCHQOMAGINFPHGIAQALWAGK 239
 DB 181 EPKPNQPRGDIILLPTAGHAIAFVQELERPELFGINPETGHEQMSNLNETQGIQAALWHKK 240
 QY 240 LFHIDLNGSGIKYDDDLRFAGDGLRAAPMLVDLLESA----GWEGPRHFDKPPRTED 294
 DB 241 LFHIDLNGQHGPKFDODLVFGHGDLLNFAFLSLVDLLENGPDGAPAYDGPGRHFDKPSRTED 300

QY 295 IDGVWASAGCMRNLYLILKERAAAFRADPEVQEAALRAARDQLAEPT--AADGLQALLAD 352
 DB 301 YDGVWESAKANIRMYLLLLKERAKAFRADPEVQEAALAAASKVAELKTPTLNPGGYAELLAD 360
 QY 353 RTAYEDEFDVA-AARGMAFERLDQLAMDHLLGAR 365
 DB 361 RSFEDYDADAVGAKGFGFVKLNQLAIEHLLGAR 394
 Search completed: March 22, 2000, 23:43:24
 Job time: 770 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2000, 23:43:59 ; Search time 39.33 Seconds
(without alignments)
680.472 Million cell updates/sec

Title: US-09-383-318-2
Perfect score: 2029
Sequence: 1 MNYQPTDRFTFGLTWGV.....GMAFERLDQLAMHLLGARG 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 59334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

SPTREMBL12:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2029	100.0	386	2 Q9ZAI3	Q9ZAI3 streptomyc
2	317	15.6	444	2 Q9X125	Q9X125 thermotoga
3	307	15.1	439	2 P77995	P77995 thermoanaer
4	303	14.9	439	2 Q9X422	Q9X422 lactococcus
5	300.5	14.8	439	2 Q9X416	Q9X416 lactococcus
6	176	8.7	169	10 Q9XFD3	Q9XFD3 oryza sativ
7	138.5	6.8	383	2 Q9X0G0	Q9X0G0 thermotoga
8	131.5	6.5	386	2 Q9XAB3	Q9XAB3 streptomyc
9	125.5	6.2	315	2 Q69950	Q69950 streptomyc
10	120	5.9	290	2 Q9Z420	Q9Z420 streptomyc
11	107	5.3	233	1 Q26349	Q26349 methanobact
12	105	5.2	397	5 Q17899	Q17899 caenorhabdi
13	105	5.2	4735	2 Q54566	Q54566 amycolatops
14	100.5	5.0	7576	2 Q9Z6A4	Q9Z6A4 streptomyc
15	97	4.8	373	2 Q9ZC00	Q9ZC00 streptomyc
16	97	4.8	501	1 Q57693	Q57693 thermoprote
17	96	4.7	415	2 Q55407	Q55407 synchocyst
18	95.5	4.7	353	4 Q43659	Q43659 homo sapien
19	95.5	4.7	367	4 Q13303	Q13303 homo sapien
20	95.5	4.7	370	2 Q9XDNO	Q9XDNO salmonella

21	95	4.7	270	2 Q9WYP7	Q9WYP7 thermotoga
22	95	4.7	293	2 Q66289	Q66289 agrobacteri
23	94.5	4.7	878	2 Q9WXH8	Q9WXH8 microblspor
24	94	4.6	287	2 P73599	P73599 synchocyst
25	94	4.6	541	2 Q50491	Q50491 streptomyc
26	94	4.6	3122	12 P89459	P89459 herpes simp
27	94	4.6	3590	2 Q43365	Q43365 bordetella
28	93.5	4.6	480	2 Q05438	Q05438 mycobacteri
29	93.5	4.6	768	2 Q086766	Q086766 streptomyc
30	93.5	4.6	1361	2 Q086617	Q086617 streptomyc
31	93	4.6	318	2 Q9WX20	Q9WX20 streptomyc
32	93	4.6	821	2 Q59241	Q59241 bacillus sp
33	92.5	4.6	367	6 Q27955	Q27955 bos taurus
34	92.5	4.6	367	11 P97381	P97381 mus musculus
35	92.5	4.6	890	2 Q53582	Q53582 streptomyc
36	92.5	4.6	1279	2 Q50314	Q50314 chlorobium
37	91.5	4.5	678	2 Q47801	Q47801 enterococcu
38	91.5	4.5	4472	2 Q33954	Q33954 streptomyc
39	91	4.5	328	1 Q9YBV2	Q9YBV2 aeropyrum p
40	91	4.5	336	2 Q06334	Q06334 mycobacteri
41	91	4.5	930	1 Q9YBL5	Q9YBL5 aeropyrum p
42	91	4.5	3729	2 Q33956	Q33956 streptomyc
43	90.5	4.5	367	11 Q64284	Q64284 mus musculus
44	90.5	4.5	451	2 Q9X8D0	Q9X8D0 streptomyc
45	90.5	4.5	798	2 Q087626	Q087626 neisseria f

ALIGNMENTS

RESULT 1

Q9ZAI3 ID Q9ZAI3 PRELIMINARY: PRT: 386 AA.
AC Q9ZAI3;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Streptomyces sp.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK;
RA BELGITH-SRIH K.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.
CC -!- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY OF THE PROTEIN.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
DR EMBL; Y15518; CAA75672.1; -.
DR HSSP; P15587; 1XYC.
DR PROSITE; PS00172; XYLOSE-ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE-ISOMERASE_2; 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.
SQ SEQUENCE 386 AA; 42702 MW; 7837ACF1 CRC32;

Query Match 100.0%; Score 2029; DB 2; Length 386;

Best Local Similarity 100.0%; Pred. No. 2.5e-143; Mismatches 0; Indels 0; Gaps 0;

Matches 386; Conservative 0;

QY 1 MNYQPTDRFTFGLTWGVQGRDPFGDTRPALDPVDVQRLAELGAYGVTFHDDLIPF 60

|||||

Db 1 MNYQPTDRFTFGLTWGVQGRDPFGDTRPALDPVDVQRLAELGAYGVTFHDDLIPF 60

QY 61 GASDTERAHVKFRQALDQATGTVPMATNLTTHPVFKAGATANDRAVRRALAKRTIR 120

|||||

Db 61 GASDTERAHVKFRQALDQATGTVPMATNLTTHPVFKAGATANDRAVRRALAKRTIR 120

QY 121 NIDLAVELGAKVIVANGGREGESGAAKOVRAALDRMKEAFDILGCVTSQGYDIRFAIE 180

|||||

Db 121 NIDLAVELGAKVIVANGGREGESGAAKOVRAALDRMKEAFDILGCVTSQGYDIRFAIE 180

RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

```
DR EMBL; AF140230; AAD30135.1; -
KW Isomerase.
FT NON_TER 1 169
FT NON_TER 169
SQ SEQUENCE 169 AA; 19511 MW; C04DD948 CRC32;

Query Match 8.7%; Score 176; DB 10; Length 169;
Best Local Similarity 30.2%; Pred. No. 4.5e-06;
Matches 42; Conservative 25; Mismatches 68; Indels 4; Gaps 1;

QY 52 FHDDLLIPFGASDTEREAH-----VKRFQALDQATGTVPMATNLFTHPVFKAGAFAND 107
   ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 27 FHDDIADPGKLTETNKNLDEIVELAKLQEBTNKPLWGTAQLFMHPYMHGAATSP 86
   ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 108 RAVRRYALRKTTRNDLAVELGAKVYVANGREGAESGAADKVRALDRMKEAFDILLGEY 167
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 VKVYAYGAQVKKALEVTHYLGGENYVFWGREGYQTLTNDMKRELDSLFLQAADVY 146
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 168 VTSQGYDIRFAIEPKPNP 166
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 147 KKKIGFNGTLLIEPKPOEP 165
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 7

```
QX0G0
ID QX0G0 PRELIMINARY; PRT; 383 AA.
AC QX0G0;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE SUGAR ISOMERASE.
GN TM1071.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 99287316.
RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.;
RT "Evidence for lateral gene transfer between Archaea and bacteria from
RT genome sequence of Thermotoga maritima."
RT Nature 399:323-329(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001767; AAD36148.1; -
SQ SEQUENCE 383 AA; 44732 MW; C0C4DEB7 CRC32;
```

Query Match 6.8%; Score 138.5; DB 2; Length 383;
Best Local Similarity 24.4%; Pred. No. 0.0089;
Matches 76; Conservative 43; Mismatches 118; Indels 75; Gaps 15;

```
QY 72 KRFQALDQATGTVPMATNLFTHPVFKAGAFANDRAVRRYALRKTIRNIDLAVELGAK 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 75 EELREFAEKGLKIGAINPFLQDDPKYKGLTNPNSEKIRKKAIAHVMCEVDIAEKTGSK 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 132 VYVANGREGAESGAADKVRALDRMKEAFDILLGEYVTSQGYDIRFAIEPKPNP---RG 188
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 VISLWLA-DGTDYPGDDPFRSRKKRLEESLRYIYENNPA---DMYLLIEYKTFEPAYHT 190
```

RESULT 8

```
QXAB3
ID QXAB3 PRELIMINARY; PRT; 386 AA.
AC QXAB3;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE PUTATIVE SUGAR ISOMERASE.
GN SCF43A.02.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SEGER K., HARRIS D.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA JAMES K.D., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE; 97000351.
RA REGENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
KW Isomerase.
SQ SEQUENCE 386 AA; 42358 MW; 99F1432A CRC32;
```

Query Match 6.5%; Score 131.5; DB 2; Length 386;
Best Local Similarity 26.0%; Pred. No. 0.03;
Matches 95; Conservative 39; Mismatches 160; Indels 71; Gaps 19;

```
QY 16 WTVGMQ-----RDFFGDATRPALDPVDVQRLAEL--GAYGVTFHDDLLIPFG 61
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 WAYNSGTRFKVFAQPGVPRDPF-----EKLD--DAAKVHEFTGAAPTVALH---IPWD 69
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 ASD--TEREAHVHVRQALDQATGTVPMATNLFTHPVFKAGAFANDRAVRRYALRXTI 119
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 RVEDYAAALAAHAEK-----RGVRIGAINSTFODDDYRLGSGICHPDAAVVRKAVDHL 122
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 120 RNIDLAVELGAKVYVANGREGAESGAADKVRALDRMKEAFDILLGEYVTSQGYDIRFAI 179
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 ECVDTMDATGSRDLKLWFA-DGTNYPGDDINSRODRLAEG---LAEYVERLGEQRMILL 178
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 180 EPKPNP---RGDILLPTIGHALAFIERLERPELYGVNPEVGHQMGALNFPHGIAQALW 236
```

```
Db 179 EYLFEPFAFTTDV--PDWGTATAC--LKLGEKAQVVDVDTGH-HAPGTNIEFIVATLLR 233
QY 237 AGKLFHIDLNGSQGIKYDQDLRFAGDLRAAFWLIV-DLLESAGWEGPRHFDFKPPRTEDI 295
Db 234 EGKLGGFDFN--SRFYADDLWGAADPFOLFIMYEVVGGGF-----TSDV 279
QY 296 DGWWSAAGCMRNLYLKERAAAFRADPEVOEALRAARL---DOLAEPTAA-DGLQALLA 351
Db 280 -----AFMLDCHNIEAKIPAIRSYVMNVQEATKALLVDGTALAEAAQAGDVLENAV 333
QY 352 DRTAY 356
Db 334 LMDAY 338

RESULT 9
ID O69950 PRELIMINARY; PRT; 315 AA.
AC O69950;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 33.5 KD PROTEIN.
GN SC3P9.10.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SEGER K.J., HARRIS D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97000351.
RA REDENBACH M., KIESER H.M., DENAPAITE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL023862; CAAL9633.1; -.
KW Hypothetical protein.
SQ SEQUENCE 315 AA; 33543 MW; AB456D17 CRC32;

Query Match 6.2%; Score 125.5; DB 2; Length 315;
Best Local Similarity 24.4%; Pred. No. 0.062;
Matches 86; Conservative 46; Mismatches 145; Indels 75; Gaps 18;

QY 21 QGRDPFGDTRPALDPV-----DVRDLAELG--AYGVTFHDDDLIPFGA 62
Db 7 QQPNAGDGG--PATHPLFGYGTNGGLADRLDDALLADLGDYGVGLTLDHMLDPLA- 64
QY 63 SDTEREAHVRFQALDQMTVPMTAT-TNLFTHPVFKAGAFAN----DRAVRRVALRK 117
Db 64 --DDLAARTRLARRLDTLGLVTGTGARYVLDPRKKGPSLLDPPQDRA--RTGL-- 119
QY 118 TIRNIDLAVELGAKVYVANGREGAESAAGKADYRAALDRKKEAFDILLGEYVTSQGDYR 177
Db 119 LLRAVDVAELGAHVHVCFSG----VTPGTDDETAWKRLAEALAPVLDAAATAG--VPL 172
QY 178 AIEPKNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNPFPHGIAQALWA 237
Db 173 AVEPEFGH-----LLATVADFHTLRGALGDPEHLGLTLDIGHCQCCLPLPPADCVRAA-A 226
QY 238 GKLFHIDLNG-QSGIKYDQDLRFAGDLRAAFWLVDLLESAGWEG-----PRHDFKPP 290
```

```
Db 227 PMLRHVOIEDMRGV--HEHLPFGDEIDFP-PVLEALAATGYQGLTVVLPDRHSHAGPH 283
QY 291 RTEDIDGVWASAGCMRNLYLKERAAAFRADPEVOEALRAARLDQLAEPSTA 342
Db 284 YAE-----RSLPFLRRRAAPPPP-----RTNRSGEPSA 313

RESULT 10
ID Q92420 PRELIMINARY; PRT; 290 AA.
AC Q92420;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE PUTATIVE ISOMERASE.
GN SCE29.10C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SAUNDERS D.C., HARRIS D.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA BENTLEY S.D., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97000351.
RA REDENBACH M., KIESER H.M., DENAPAITE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL035707; CAB38885.1; -.
KW Isomerase.
SQ SEQUENCE 290 AA; 31122 MW; D8096EF4 CRC32;

Query Match 5.9%; Score 120; DB 2; Length 290;
Best Local Similarity 23.9%; Pred. No. 0.14;
Matches 72; Conservative 48; Mismatches 133; Indels 48; Gaps 17;

QY 22 GRDPFGDTRPALDPVDVORLAEIG--AYGVTFHDDDLIPFGASDTEREAHVRFQALD 79
Db 10 GTNGFGD--HPLEDALAV--LADLGYEGVGLTLDPRHLDPF--ADDLPH--RLRRLAARD 62
QY 80 ATGEMTVPMATTNLT-----HPVEKAGAFANDRAVRYALRKTRINDLAVELGAKV 132
Db 63 RUGLAVVVTGGRYVLDPRKHPQVILMSA-----EGAGRRVDL--LURAVRIAADLGAEE 115
QY 133 YVANGREGAESAAGKADYRAALDRKKEAFDILLGEYVTSQYDIRFAIEPKNEPRGDI 192
Db 116 VSPWSGAAPADTPR---QVWDRLLAGCGTVVEAARAGVVLGF--EPEPG-----MFV 164
QY 193 PTGHALAFIERLERPELYGVNPEVGEQAGLNPFPHGIAQAL--WAGKLFHIDLNGSQG 250
Db 165 DTLDAYDELCCRLLGGPHPLGLTLDIGHRCLE---POPVADCVRRVADRLVNVOIEDMR 221
QY 251 IKYDQDLRFAGDLRAAFWLVDLLESAGWEG-----PRHDFKPP-RTEDIDGVWASAA 303
Db 222 GTHEH-LEFGSGEIDFP-PVLSALAATGYRGLVSVLPDRHSHAAPVARRSLDFLRAAD 279
QY 304 G 304
Db 280 G 280
```

RESULT 11
 Q26349 ID O26349 PRELIMINARY; PRT; 233 AA.
 AC O26349
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)
 DE CONSERVED PROTEIN.
 GN MTH247
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H:
 RX MEDLINE; 98037514.
 RA SMITH D.R., DOUGHTY-SPAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
 RA ALDREDGE T., BASHIRADEH R., BLAKELY D., COOK R., GILBERT K.,
 RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTIER B., QIU D.,
 RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
 RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
 RA MCOUGALL S., SHMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,
 RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.,
 RA *Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.*;
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL; AE000811; AAB84753.1; F9D4C1C3 CRC32;
 SQ SEQUENCE 233 AA; 25067 MW; F9D4C1C3 CRC32;

Query Match 5.3%; Score 107; DB 1; Length 233;
 Best Local Similarity 26.6%; Pred. No. 0.96;
 Matches 54; Conservative 32; Mismatches 83; Indels 34; Gaps 10;
 QY 96 PVFKAGATANDRAVRRVALRKTIRNIDLAVELGAKYVWGGREGAAGAAKDVRAALD 155
 DB 57 PISDINIASLND-TYREASIREVASVDLAVELSDVVVHPGTPFLGRPYRDLIA--E 113
 QY 156 RMKEAFDILLGELVTSQGYDIRFAIPKPNPR--GDILLPTIGHALAFIERLERELXGV 213
 DB 114 RNLESIAISEVASDRG-----GVVPE-NMPLEGPLL-----RELELWGV 155
 QY 214 NPEVGEQAGLNFPHGI-----AQAALWAGKLFHIDLNGSQGIKYDQD-LRFAGAGDLRAA 267
 DB 156 AEELG--VNVTLDAHAATMGVSTEEVSAHGVHLSLSDTGEVDSHDALGSGSLDFQA- 213
 QY 268 FWLVLLSAGWEGPRHDFKPP 290
 DB 213 --LLDGLRGAGYTGVLTVETKTP 233

RESULT 12
 Q17899 ID Q17899 PRELIMINARY; PRT; 397 AA.
 AC Q17899
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE C10C5.4 PROTEIN.
 GN C10C5.4
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MATTHEWS P.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z68214; CAA92447.1; -;
 DR PFAM; PF01546; Peptidase_M20; 1.
 SQ SEQUENCE 397 AA; 45242 MW; 0963D445 CRC32;

Query Match 5.2%; Score 105; DB 5; Length 397;

Best Local Similarity 20.5%; Pred. No. 2.9;
 Matches 68; Conservative 46; Mismatches 115; Indels 102; Gaps 18;
 QY 83 MTVP-----MATTNLFTH----PVFKA-----GATANDRAVRRVALRKT-----IRNI 122
 DB 58 MTIFGSELPESIMLYSHTDVVPTFREHWTHTDPSAFKEDGNIFARGAQDTKCLGVQHV 117
 QY 123 DLAVELGAK-----VYVWANGREGAESAAGKDVRAALD-----RMKEAFDILLGELV 169
 DB 118 EAFNLFAKAGAKQWKRTIHVWGPDE--ETGHINGMKGFVETNEFKLINIGFSLDEGHS 175
 QY 170 SOGYDIRFAIE-----PKPNEPRGDILLPTIGHALAFIE-----R 204
 DB 176 ENGVLTFYAEVTVWLKVTAPGNP-----GHGSQFMENTAMEKIERFLASARA 226
 QY 205 LERELGVNP--EVGHEOMAGLNFPHGIAQALWAGKLF--HIDLNGSQGIKYDQDLRF 260
 DB 227 NEQKELLEKNFTWALGDVTTLNVLNKGQVFNVIPEKFEAYVDIR-----LTPNODF--- 280
 QY 261 AGDLRAAF--WLVDLLESAGWEGPRHDFKPP--PRTEIDIGWAS----- 302
 DB 280 -GEIRAKLDQWVKDAGSGVTFEFSQHSYDKPVSPHTRD-DFWAAAFEDSLKQENCEFTTE 337
 QY 302 -AAGCMRNLYLKERAAAFRADPEVQEAALRA 331
 DB 338 VCVGSTDSRFVRKAGVPAINFSPMINTPVA 368

RESULT 13
 O54666 ID O54666 PRELIMINARY; PRT; 4735 AA.
 AC O54666;
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE POLYKETIDE SYNTHASE.
 GN RIFA.
 OS Amycolatopsis mediterranei.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S699;
 RA KIM C.G., YU T.W., FRYHLE C., HANDA S., FLOSS H.G.;
 RL J. Biol. Chem. 0:0-0(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S699;
 RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., YU T.W.,
 RA TAYLOR M., HOFFMANN D., KIM C.G., ZHANG X., HUTCHINSON C.R.,
 RA FLOSS H.G.;
 RL Chem. Biol. 5:0-0(0002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S699;
 RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., HUTCHINSON C.R.,
 RA FLOSS H.G.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LBG A3136;
 RA SCHUPP T., TOUPET C., ENGEL N., GOFF S.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF040570; AAC01710.1; -;
 DR EMBL; AJ223012; CAAL1035.1; -;
 DR HSSP; P08659; ILIC.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PFAM; PF00550; pp-binding; 4.
 DR PFAM; PF00109; ketoacyl-synt; 3.
 DR PFAM; PF00698; Acyl_transf; 3.
 DR PFAM; PF00501; AMP-binding; 1.
 DR PRINTS; PR00154; AMPBINDING.

KW Transferase. 4735 AA; 495213 MW; FC634DAE CRC32;

Query Match 5.2%; Score 105; DB 2; Length 4735;
Best Local Similarity 24.5%; Pred. No. 1,1e-02;
Matches 92; Conservative 46; Mismatches 140; Indels 98; Gaps 22;

QY 3 YQPTPEDRTFF-----GLTVGQGRDPFGDTRPALDP-VDVQRLAEIAGVGVTFHDD 56
DB 4417 FEDTPLDAFVFSIAGGGGQ--PAG-RANAVLDALVEWRRARGLKATSIANGALD 4473
QY 57 LIPFGASDTEREAHYKRFQALDAGTMTVPMTNLTTHPVKAGAFANDRAVRYALR 116
DB 4474 QIGIGMD-----EAALAQLRRR-GVIMPAPPLAVTAM-----VQAVAGNEKAV----- 4516
QY 117 KTRINIDLAVELGAKVYVANGREG---AESGAAKDV-RAALDRMKEAFDILLGEYVTSQ 172
DB 4516 -AVADMWRAFIPTSV-----RPSPLFADLPEAKAILRAAQDDGD-----GDTASSLA 4565
QY 173 YDIRFAIEPKPNRPGDILLPTI-GHALAFIERLERPELYGVNP-----EVGHEQMAGLN 226
DB 4566 DSLRAV-----PDAEQNRILLKLVRGHASTV---LHSGAEGIGPRQAFQEVGFDSLAAYN 4618
QY 227 FPHGTAQALWACKLPHIDLNGOSGKIYDODLRF-----AGDLRAAFWLYDLESAA-- 278
DB 4619 LRN-----SUHAATGLRLPATLIFDYPTPEALVGYLR-----VELLREADD 4659
QY 278 GWGPRHFDKPPRTEDIDGVNAAA-----AGCMRNLYLKERAAAFRAADPEVEALR 330
DB 4660 GLDG-----REDDLRLVLAANVFAFKEAGVLTLLGLADITGTEGTAETEARP 4710
QY 331 AARLDQAEPTAADGL 346
DB 4711 AADDAELIDALDISGL 4726

RESULT 14

ID Q9ZGA4 PRELIMINARY; PRT; 7576 AA.
AC Q9ZGA4;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE FK506 POLYKETIDE SYNTHASE.
GN FKBB.
OS Streptomyces sp. MA6548.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MA6548;
RX MEDLINE: 98451508.
RA MOTAMEDJ H., SHAFIEE A.;
RT "The biosynthetic gene cluster for the macrolactone ring of the
immunosuppressant FK506.";
RL Eur. J. Biochem. 256:528-534(1998).
DR EMBL: AF082100; AAC68815.1; -.
DR HSSP: P08659; 1LCI.
DR PROSITE: PS00455; AMP_BINDING: 1.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE: 4.
KW Transferase.
SQ SEQUENCE 7576 AA; 790105 MW; 42E87D19 CRC32;

Query Match 5.0%; Score 100.5; DB 2; Length 7576;
Best Local Similarity 24.7%; Pred. No. 4.5e-02;
Matches 92; Conservative 31; Mismatches 167; Indels 83; Gaps 19;

QY 25 PFGDTRPALDPVQRLAEIAGVGVTFHDDLLIPFGASDTEREAHVKFRQALDAGTMT 84
DB 3203 PAAPAPLPADPVP-----EDVVVPSVARTPDGVRDV-----AGRLTALT 3242

QY 85 VPMA-----TTNLFTH-----PVFKAGAFANDR---AVRRYALRKTRINIDLAVEIG 129
DB 3243 APAAAIGHSLATRTTAMRRRAVVPARDAEAFARGEVEPVGVRGTADVTDTTRAVFVPPGOG 3302
QY 130 AKVYVANGREGAESAAGKDVRAALDRMKEAFDILLGEYVTSQYDIRFAIEPKPNP--- 187
DB 3303 SQ-----WDGM-CAELLATEPVFAR--RLGECALAPYT--GWDLDDVIARRPGAPELD 3352
QY 187 RGDILLPTIGHALAFIERLERPELYGVNPE--VGHEQ-----MAG-LNFPHGIAQALW 236
DB 3353 RVDVYQVPASFAMVALAELWR--AHGVAPAAVVGHSQGEVAAAACVAGVLTLDAAKVVWAL 3410
QY 237 AGKLFHIDLNGOSGKIYDODLRFAGAGDLRAAFWLYDLESAGWEGPRHFDKPRPTEDID 296
DB 3411 RSRVLATERAGHGM-----VSVPPADFADAAW-AGRLEVAAVNGPAST----- 3454
QY 297 GWASAAGCMRNLYLKERAAAFRAADPEVQ-----PALRAARLDQAEPT-AADGLQALLA 351
DB 3454 -VVAGNADAVEELLAAATPHARRIANDYASHTAHVESIRGALLDALADLTTPGAPEIPFFST 3512
QY 352 DRTAYEDEFVDAA 364
DB 3513 VDEAWLDRPADAA 3525

RESULT 15

ID Q9ZC00 PRELIMINARY; PRT; 373 AA.
AC Q9ZC00;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE HYPOTHETICAL 40.7 KD PROTEIN.
GN SC1B6.22C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SEGER K.J., HARRIS D.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE: 97000351.
RA REDENBACH M., KIESER H.M., DENAPATTE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL033505; CAA22050.1; -.
KW Hypothetical protein.
SQ SEQUENCE 373 AA; 40702 MW; AEA38C6C CRC32;

Query Match 4.8%; Score 97; DB 2; Length 373;
Best Local Similarity 22.4%; Pred. No. 10;
Matches 72; Conservative 29; Mismatches 108; Indels 112; Gaps 17;

QY 93 FTHPVFK-----AGAFANDRAVRYALRKTRINIDLAVELGAKVYVANGREGAEGAA 147
DB 5 FLHPLYEHOGPNASVYVDLSRHTEDTPHERELTAATAAVAREL-----ASEGADATC 55
QY 148 KDVRAALDRMKEAFDILLGEYVTSQYDIRFAIEPKPNRPGDILLPTIGHALAFIERLER 207
DB 56 RAVREAVDELRAATD-----PHGRALFACAGQ-VVLDPALAR 91

```

Qy 208 PELXGVNPEVGEQAGLNFPHGIAQALWAGKLFHI-----DLNGOSGI-----KYD 254
Db 92 P-----PYGGTTADWA-PLPHVTPLLDLAGEDPVCVVAYIDRKGAD 131
Qy 255 QDLRFQ-----AGDLRAAFNLVDLLESAGWEGPRHFDFKPPRTEDIDGVNA-----S 301
Db 132 FELRSALGSSDAGGVTRQWPVHRTSSADW-SERHFQLR-----VENTWEHNAETADA 184
Qy 302 AAGCMR-----NYLIL-----KERAAA-----FRADPEVCEALRAA---RLDQLAEPTRAADG 345
Db 185 LAVCOETGADLLILVGDRERRSVHERLPLRLQERVAEASRGAGSRLLDDEVEGLRDDH 244
Qy 346 LQALLADRTAYEDEFVDAAAR 366
Db 245 VRA-----RAREDLDRFLAAR 260

```

Search completed: March 22, 2000, 23:47:30
Job time: 211 sec

GenCore version 4.5
 Copyright (c) 1993 - 2000 CompuGen Ltd.
 OM protein - protein search, using sw model
 Run on: March 22, 2000, 23:43:29 ; Search time 27.27 Seconds
 (without alignments)
 422.730 Million cell updates/sec
 Title: US-09-383-318-2
 Perfect score: 2029
 Sequence: 1 MNYQPTEDRTFGLTWGV.....GMAFERLDQLAMHLLGARG 386
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 82229 seqs, 29864866 residues
 Total number of hits satisfying chosen parameters: 82229
 Minimum DB seq length: 0
 Maximum DB seq length: 1000000
 Post-processing: Minimum Match 0%
 Listing first 45 summaries
 Database : SwissProt_38:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1916	94.4	386	XYLA_STROL	P15587 streptomyc
2	1900	93.6	388	XYLA_STRM	P37031 streptomyc
3	1887	93.0	388	XYLA_STRVO	P09033 streptomyc
4	1872	92.3	387	XYLA_STRRU	P24300 streptomyc
5	1863	91.8	390	XYLA_STRAL	P24299 streptomyc
6	1834	90.4	388	XYLA_STRDI	P05910 streptomyc
7	1534	75.6	393	XYLA_STRRO	P22857 streptomyc
8	1346	66.4	393	XYLA_ACTMI	P12851 actinoplan
9	1344	66.3	394	XYLA_AMSP	P10854 ampullari
10	1277	62.9	394	XYLA_ARTS7	P12070 arthroba
11	1127	55.5	387	XYLA_THETH	P26997 thermus aqu
12	1100	54.2	387	XYLA_THECA	P56681 thermus aqu
13	333.5	16.4	449	XYLA_LACPE	P21938 lactobacill
14	329	16.2	448	XYLA_BACLI	P77832 bacillus li
15	324.5	16.0	435	XYLA_TETHA	O82845 tetragenoco
16	320	15.8	444	XYLA_THENE	P45887 thermotoga
17	318	15.7	439	XYLA_CLOTS	P29441 clostridium
18	314	15.5	449	XYLA_LACBR	P29443 lactobacill
19	312.5	15.4	439	XYLA_STAXY	P27157 staphylococ
20	310.5	15.3	439	XYLA_THETU	P19148 thermoanaer
21	310	15.3	439	XYLA_THESA	P30435 thermoanaer
22	301.5	14.9	438	XYLA_THETT	P22842 thermoanaer
23	299.5	14.8	445	XYLA_BACME	O08325 bacillus me
24	299	14.7	441	XYLA_BACST	P54272 bacillus st
25	299	14.7	441	XYLA_BACST	P54273 bacillus st
26	297.5	14.7	445	XYLA_BACSU	P04788 bacillus su
27	277.5	13.7	440	XYLA_ECOLI	P00944 escherichia
28	273.5	13.5	440	XYLA_KLEPN	P29442 klebsiella
29	244.5	12.1	479	XYLA_HAEVU	Q40082 hordeum vul
30	239.5	11.8	439	XYLA_HAEIN	P44398 haemophilus
31	106	5.2	3491	ERYL_SACER	Q03131 saccharopol
32	98.5	4.9	1061	TRC4_ECOLI	P27189 escherichia
33	97	4.8	1381	VCAP_EBV	P03226 Epstein-bar
34	96	4.7	800	GUN_BACSI	P06564 bacillus sp

ALIGNMENTS

RESULT 1	XYLA_STROL	STANDARD;	PRT;	386 AA.
35	94	4.6	3591	1 FHAB_BORPE
36	93.5	4.6	1403	1 VG22_HSVII
37	93	4.6	540	1 CH62_MYCLE
38	92.5	4.6	641	1 HRPK_PSESY
39	92.5	4.6	1486	1 MUKB_ECOLI
40	92	4.5	539	1 CH62_MYCTU
41	91.5	4.5	392	1 NOLC_RHIFR
42	91.5	4.5	482	1 NODT_RHILV
43	91	4.5	293	1 VIB9_AGRTR
44	91	4.5	389	1 YS96_MYCTU
45	90.5	4.5	3567	1 ERY2_SACER
SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.81 ANGSTROMS).				
RP	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.81 ANGSTROMS).			
RX	MEDLINE: 94235606.			
RA	LAVIE A., ALLEN K., PETSKO G.A., RINGE D.;			
RT	"X-ray crystallographic structures of D-xylose isomerase-substrate complexes position the substrate and provide evidence for metal movement during catalysis.";			
RL	Biochemistry 33:5469-5480(1994).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).			
RX	MEDLINE: 90057422.			
RA	FARBER G.K., GLASFELD A., TIRABY G., RINGE D., PETSKO G.A.;			
RT	"Crystallographic studies of the mechanism of xylose isomerase.";			
RN	Biochemistry 28:7289-7297(1989).			
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).			
RX	MEDLINE: 89220983.			
RA	FARBER G.K., PETSKO G.A., RINGE D.;			
RT	"The 3.0 A crystal structure of xylose isomerase from Streptomyces olivochromogenes";			
RL	Protein Eng. 1:459-466(1987).			
CC	- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.			
CC	- CATALYTIC ACTIVITY: D-XYLOSE - D-XYLULOSE.			
CC	- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.			
CC	- SUBUNIT: HOMOTETRAMER.			
CC	- SUBCELLULAR LOCATION: CYTOPLASMIC.			
CC	- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.			
DR	PIR; S28986; S28986.			
DR	PDB; 1XYA; 31-MAY-94.			
DR	PDB; 1XYB; 31-MAY-94.			
DR	PDB; 1XYC; 31-MAY-94.			
DR	PDB; 1XYL; 31-MAY-94.			
DR	PDB; 1XYM; 31-MAY-94.			
DR	PDB; 2GYI; 10-JUL-95.			
DR	PROSITE; PS00172; XYLOSE_ISOMERASE_1; 1.			
DR	PROSITE; PS00173; XYLOSE_ISOMERASE_2; 1.			
DR	PFAM; PF00259; Xylose_isom; 1.			
KW	Isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.			
FT	INIT_MET 0			
FT	ACT_SITE 53 53 BY SIMILARITY.			
FT	METAL 180 180 MAGNESIUM (BY SIMILARITY).			
FT	METAL 216 216 MAGNESIUM (BY SIMILARITY).			
FT	METAL 244 244 MAGNESIUM (BY SIMILARITY).			
FT	METAL 286 286 MAGNESIUM (BY SIMILARITY).			
FT	STRAND 12 13			

P12255 bordetella
 Q00105 ictaluriid h
 P09239 mycobacteri
 P41501 pseudomonas
 P22523 escherichia
 P06806 mycobacteri
 P26508 rhizobium f
 P15727 rhizobium l
 P05358 agrobacteri
 Q10817 mycobacteri
 Q03132 saccharopol

```
FT TURN 22 25
FT HELIX 36 40
FT TURN 41 42
FT STRAND 52 52
FT TURN 55 57
FT TURN 64 71
FT TURN 75 76
FT STRAND 84 84
FT TURN 101 102
FT TURN 104 105
FT HELIX 106 120
FT STRAND 125 127
FT TURN 130 131
FT TURN 141 142
FT TURN 144 149
FT HELIX 150 153
FT TURN 154 155
FT STRAND 159 160
FT STRAND 177 179
FT TURN 197 198
FT HELIX 199 205
FT TURN 206 206
FT TURN 221 222
FT HELIX 233 239
FT STRAND 247 247
FT TURN 260 262
FT TURN 267 275
FT STRAND 283 284
FT TURN 295 296
FT TURN 298 300
FT HELIX 302 305
FT TURN 306 307
FT HELIX 308 311
FT TURN 312 320
FT HELIX 322 326
FT TURN 327 328
FT HELIX 343 345
FT TURN 346 347
FT TURN 362 363
FT HELIX 364 367
FT TURN 368 369
SQ SEQUENCE 386 AA; 42792 MW; 0F0B8308 CRC32;

Query Match 94.4%; Score 1916; DB 1; Length 386;
Best Local Similarity 94.3%; Pred. No. 6.1e-143;
Matches 364; Conservative 13; Mismatches 7; Indels 2; Gaps 2;

QY 2 NYQTPEDRFTGLTVGWQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIPF 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 SQQTPEDRFTGLTVGWQGRDPFGDTRPALDPVETVQRLAELGAGVTFHDDDLIPF 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 GASDTEREAVKRFQALDQATGMTVPMATNLTTHPVFKAGAFANDRAVRVYALRKTTIR 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 GSSDTERESHKRFQALDQATGMTVPMATNLTTHPVFKDGGFTANDRDVRYALRKTTIR 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 NIDLAVELGAKYVAVWGREGAESGAADKDVRAALDRMKEAFDLGELYTSQGYDIRFATE 180
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 NIDLAVELGAKYVAVWGREGAESGAADKDVRAALDRMKEAFDLGELYTSQGYDIRFATE 180
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 PKPNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGKL 240
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 PKPNEPRGDILLPTVGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGKL 240
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 241 FHIIDNGQSGIKYDQDLRFAGDGLRAAFWLVDLLSAGWEGPRHFDKPPRTEDIDGVWA 300
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 FHIIDNGQSGIKYDQDLRFAGDGLRAAFWLVDLLSAGWEGPRHFDKPPRTEDIDGVWA 300
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 301 SAAGCMRNVLILKRAAFAADPEVQEAALRAARLDQAEPTAAGDLQALLADRTAYEDFD 360
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 SAAGCMRNVLILKRAAFAADPEVQEAALRAARLDQAEPTAAGDLQALLADRTAYEDFD 360
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 361 VD-AAARGMAFERLDQLAMDHLLGAR 385
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Db 361 VDAARGMAFERLDQLAMDHLLGAR 386
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 2
XYLA_STRMR
ID XYLA_STRMR STANDARD; PRT; 388 AA.
AC P37031;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Streptomyces murinus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RA RASMUSSEN H., LA COUR T., NYBORG J., SCHUELEIN M.;
RT "Structure determination of glucose isomerase from Streptomyces
   murinus at 2.6-A resolution."
RL Acta Crystallogr. D 50:124-131(1994).
CC -|- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
CC -|- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.
CC -|- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC -|- SUBUNIT: HOMOTETRAMER.
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -|- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AL0243; CAA00885.1; -
PDB; 1DXI; 31-JAN-94.
DR PROSITE; PS00172; XYLOSE_ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; 1.
PFAM; PF00259; Xylose_isom; 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.
FT ACT_SITE 54 54 BY SIMILARITY.
FT METAL 181 181 MAGNESIUM.
FT METAL 217 217 MAGNESIUM.
FT METAL 245 245 MAGNESIUM.
FT METAL 287 287 MAGNESIUM.
FT TURN 7 8
FT STRAND 11 14
FT HELIX 15 18
FT TURN 19 19
FT HELIX 36 46
FT TURN 47 47
FT STRAND 50 54
FT HELIX 55 58
FT HELIX 65 82
FT TURN 83 83
FT STRAND 85 85
FT STRAND 88 90
FT TURN 97 99
FT TURN 104 105
FT HELIX 109 127
FT TURN 128 130
FT STRAND 133 136
FT STRAND 142 143
FT HELIX 151 171
FT TURN 172 173
FT STRAND 178 180
FT STRAND 190 191
FT HELIX 196 203
FT TURN 204 205
FT TURN 209 211
FT HELIX 211
```

FT TURN 212 212
 FT STRAND 214 214
 FT STRAND 217 217
 FT HELIX 218 222
 FT TURN 223 225
 FT HELIX 228 235
 FT TURN 236 239
 FT STRAND 244 246
 FT STRAND 248 248
 FT STRAND 258 258
 FT TURN 260 261
 FT HELIX 265 277
 FT TURN 278 279
 FT STRAND 284 286
 FT TURN 292 293
 FT HELIX 296 320
 FT TURN 321 322
 FT HELIX 324 332
 FT TURN 333 338
 FT HELIX 335 338
 FT TURN 339 339
 FT TURN 349 350
 FT TURN 354 359
 FT HELIX 362 366
 FT TURN 367 367
 FT HELIX 372 378
 FT TURN 379 380
 SQ SEQUENCE 388 AA; 42772 MW; F62EE677 CRC32;

Query Match 93.6%; Score 1900; DB 1; Length 388;
 Best Local Similarity 92.8%; Pred. No. 1.le-141;
 Matches 360; Conservative 14; Mismatches 12; Indels 2; Gaps 2;

QY 1 MNYQPTPEDFTGLTWVGWQGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
 Db 1 MSQPTPEDFTGLTWVGWQGRDPFGDTRPALDPVDVETVORLAELGAYGVTFHDDDLIP 60
 QY 60 FGASDTEREAHVKFRQALDAGMTVPMTATNLTFTHPVFKAGFTANDRAVRVYALRKT 119
 Db 61 FGSDTERESHKFRQALDAGMTVPMTATNLTFTHPVFKDGGFTANDRDVRYALRKT 120
 QY 120 RNTDLAVELGAKVYVANGGREGAEGSAKDVRAALDRMKEAFDLLGGEYVTSQGYDIRFAI 179
 Db 121 GNTDLAAELGAKTYVANGGREGAEGSAKDVRAALDRMKEAFDLLGGEYVTAQGYDLRFAI 180
 QY 180 EPKPNPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 239
 Db 181 EPKPNPRGDILLPTVGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 240
 QY 240 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 299
 Db 241 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 300
 QY 300 ASAGCMRNLYILKERAARADPEVQEAALRAALDQAEPTADGLQALLADRTAYEDF 359
 Db 301 ASAGCMRNLYILKERAARADPEVQEAALRAALDQAEPTADGLQALLADRTAYEDF 360
 QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
 Db 361 DVDAARGMAFERLDQALMDHLLGARG 388

RESULT 3
 XYLX_STRVO STANDARD; PRT; 388 AA.
 ID XYLX_STRVO
 AC P09033;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE XYLOSE ISOMERASE (EC 5.3.1.5).
 GN XYLX.
 OS Streptomyces violaceoniger.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
 RX STRAIN-CBS 409-73;
 MEDLINE; 89016633.
 RA DROCOURT D., BEJAR S., CALMELS T., REYNES J.-P., TIRABY G.;
 RT "Nucleotide sequence of the xylose isomerase gene from Streptomyces
 violaceoniger";
 RL Nucleic Acids Res. 16:9337-9337(1988).
 CC - FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
 CC - CATALYTIC ACTIVITY: D-XYLOSE = D-XYULOSE.
 CC - COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
 CC - SUBUNIT: HOMOTETRAMER.
 CC - SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC - SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).
 DR EMBL; X12816; CAA31304.1; -;
 DR EMBL; M36269; AAA26839.1; -;
 DR PIR; S01436; ISSMXV.
 DR HSP; P37031; 1DXI.
 DR PROSITE; PS00172; XYLOSE_ISOMERASE_1; 1.
 DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; 1.
 DR FRAM; PF00359; Xylose_isom; 1.
 KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.
 FT INIT MET 0 0 BY SIMILARITY.
 FT ACT_SITE 53 53 MAGNESIUM (BY SIMILARITY).
 FT METAL 180 180 MAGNESIUM (BY SIMILARITY).
 FT METAL 216 216 MAGNESIUM (BY SIMILARITY).
 FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
 FT METAL 286 286 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 388 AA; 43013 MW; 37D69108 CRC32;

Query Match 93.0%; Score 1887.5; DB 1; Length 388;
 Best Local Similarity 92.3%; Pred. No. 1e-140;
 Matches 358; Conservative 15; Mismatches 12; Indels 3; Gaps 2;

QY 2 NYQPTPEDFTGLTWVGWQGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 60
 Db 1 SFQPTPEDFTGLTWVGWQGRDPFGDTRPALDPVDVETVORLAELGAYGVTFHDDDLIP 60
 QY 61 GASDTEREAHVKFRQALDAGMTVPMTATNLTFTHPVFKAGFTANDRAVRVYALRKTIR 120
 Db 61 GSSDTERESHKFRQALDAGMTVPMTATNLTFTHPVFKDGGFTANDRDVRYALRKTIR 120
 QY 121 NIDLAVELGAKVYVANGGREGAEGSAKDVRAALDRMKEAFDLLGGEYVTSQGYDIRFAIE 180
 Db 121 NIDLAAELGAKTYVANGGREGAEGSAKDVRAALDRMKEAFDLLGGEYVTAQGYDLRFAIE 180
 QY 181 PKPNPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGKL 240
 Db 181 PKPNPRGDILLPTVGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGKL 240
 QY 241 FHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVWA 300
 Db 241 FHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVWA 300
 QY 301 SAAGCMRNLYILKERAARADPEVQEAALRAALDQAEPTADGLQALLADRTAYEDFD 360
 Db 301 SAAGCMRNLYILKERAARADPEVQEAALRAALDQAEPTADGLQALLADRTAYEDFD 360
 QY 361 VDAARGMA--FERLDQALMDHLLGARG 386
 Db 361 VEAARGMA--FERLDQALMDHLLGARG 388


```

Db 61 GSSDSEEHVKRFRQALDDTGKVPMTNLTTHPVFKDGGFTANDRDRVRRYALRKRTIR 120
QY 121 NIDLAVELGAKVYVANGREGAEGSAKADYRAALDRMKEAFDGLLGEYVTSQGYDIRFAIE 180
Db 121 NIDLAVELGAETVANGREGAEGSAKADYRDALDRMKEAFDGLLGEYVTSQGYDIRFAIE 180
QY 181 PKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPPHGIQAALWAGKL 240
Db 181 PKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPPHGIQAALWAGKL 240
QY 241 FHIDLNGOSGIKVDQDLRFAGAGDLRAAFWLVLDLLESAGWEGPRHDFKPPRTEDIDGVWA 300
Db 241 FHIDLNGQGIKVDQDLRFAGAGDLRAAFWLVLDLLESAGISGPRHDFKPPRTEDIDGVWA 300
QY 301 SAAGCMRNILKERRAAFRADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDFD 360
Db 301 SAAGCMRNILKERRAAFRADPEVQEAALRASRLDELARPTAADGLQALLADRSAPFEFD 360
QY 361 VD-AAARGMAFERLDQANDHLLGARG 386
Db 361 VDAARGMAFERLDQANDHLLGARG 387

RESULT 5
XYLA_STRAL STANDARD; PRT; 390 AA.
AC P24299;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Streptomyces albus 9.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1] X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RX MEDLINE; 91197439.
RA DAUTER Z., TERRY H., WITZEL H., WILSON K.S.;
RT "Refinement of glucose isomerase from Streptomyces albus at 1.65 A
RT with data from an imaging plate."
RL Acta Crystallogr. B 46:833-845(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE; 89211422.
RA DAUTER Z., DAUTER M., HEMKER J., WITZEL H., WILSON K.S.;
RT "Crystallisation and preliminary analysis of glucose isomerase from
RT Streptomyces albus."
RL FEBS Lett. 247:1-8(1989).
CC -!- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
CC -!- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.
CC -!- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
DR PDB; 6XLA; 15-JUL-92.
DR PROSITE; PS00172; XYLOSE ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE ISOMERASE_2; 1.
DR PFAM; PF00359; Xylose_isom; 1.
KW isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.
FT ACT_SITE 53 53 BY SIMILARITY.
FT METAL 180 180 MAGNESIUM (BY SIMILARITY).
FT METAL 216 216 MAGNESIUM (BY SIMILARITY).
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 286 286 MAGNESIUM (BY SIMILARITY).
FT HELIX 6 8
FT STRAND 10 13
FT HELIX 14 17
FT TURN 18 18
FT STRAND 23 23
FT TURN 24 25
FT STRAND 26 26

```

```

-FT HELIX 35 45
-FT TURN 46 46
-FT STRAND 49 53
-FT HELIX 54 57
-FT TURN 60 61
-FT HELIX 64 81
-FT TURN 82 82
-FT STRAND 84 84
-FT STRAND 87 89
-FT HELIX 96 98
-FT TURN 99 100
-FT TURN 103 104
-FT HELIX 108 127
-FT TURN 128 129
-FT STRAND 132 135
-FT TURN 138 139
-FT STRAND 141 142
-FT HELIX 145 147
-FT HELIX 150 171
-FT TURN 172 172
-FT STRAND 176 179
-FT STRAND 189 190
-FT HELIX 195 202
-FT TURN 203 204
-FT HELIX 208 210
-FT STRAND 211 213
-FT STRAND 216 216
-FT HELIX 217 221
-FT TURN 222 224
-FT HELIX 227 236
-FT TURN 237 238
-FT STRAND 240 245
-FT STRAND 244 245
-FT STRAND 247 247
-FT STRAND 257 257
-FT TURN 259 260
-FT HELIX 264 276
-FT TURN 277 278
-FT STRAND 283 285
-FT TURN 291 292
-FT HELIX 295 321
-FT HELIX 323 331
-FT TURN 332 333
-FT HELIX 334 337
-FT TURN 338 338
-FT TURN 343 344
-FT HELIX 346 351
-FT HELIX 353 355
-FT TURN 356 358
-FT HELIX 361 366
-FT HELIX 371 383
-FT TURN 384 384
SQ SEQUENCE 390 AA; 43289 MW; 88ABE836 CRC32;

Query Match 91.8%; Score 1863; DB 1; Length 390;
Best Local Similarity 91.7%; Pred. No. 8.7e-139;
Matches 355; Conservative 15; Mismatches 15; Indels 2; Gaps 2;

QY 2 NYOPTPEDRFTFGLWTVGNQGRDPFGDATERPALDPVD-VQRLAELGAYGVTFHDDDLIPF 60
Db 1 NYOPTPEDRFTFGLWTVGNQGRDPFGDATERPALDPVD-VQRLAELGAYGVTFHDDDLIPF 60
QY 61 GASDTEREAHVKKRFRQALDDTGKVPMTNLTTHPVFKAGAFANDRAVRRYALRKRTIR 120
Db 61 GSSDSERYEHVKRFRQALDDTGKVPMTNLTTHPVFKDGGFTANDRDRVRRYALRKRTIR 120
QY 121 NIDLAVELGAKVYVANGREGAEGSAKADYRAALDRMKEAFDGLLGEYVTSQGYDIRFAIE 180
Db 121 NIDLAVELGAETVANGREGAEGSAKADYRDALDRMKEAFDGLLGEYVTSQGYDIRFAIE 180
QY 181 PKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPPHGIQAALWAGKL 240

```

Db 181 PKPNEPRGDILLPTTGHAFIERLERPELYGVNPEVGHQMAGLNPHGIAQALWAGK 240
 QY 241 FHIDLNGSGIKYDODLRFGAGDLRAAFWLVLLSAGWEGPRHDFKPPRTEDIDGVA 300
 Db 241 FHIDLNGSGIKYDODLRFGAGDLRAAFWLVLLSAGWEGPRHDFKPPRTEDIDGVA 300
 QY 301 SAAGCMRNLYLILKRAAFAFRADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDFD 360
 Db 301 SAAGCMRNLYLILKRAAFAFRADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDFD 360
 QY 361 VD-AAARGMAFERLDQALDHLGARG 386
 Db 361 VDAAAARGMAFERLDQALDHLGARG 387

RESULT 6
 XYL_A_STRDI STANDARD; PRT; 388 AA.
 AC P50910;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE XYLASE ISOMERASE (EC 5.3.1.5).
 GN XYL.
 OS Streptomyces diastaticus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NO. 7 / M1033;
 RX MEDLINE; 95101954.
 RA WANG Y., HUANG Z., DAI X., LIU J., CUI T., NIU L., WANG C., XU X.;
 RT "The sequence of xylose isomerase gene from Streptomyces diastaticus No. 7 M1033".
 RL Chin. J. Biotechnol. 10:97-103(1994).
 CC -!- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
 CC -!- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.
 CC -!- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
 CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SIMILARITY: BELONGS TO THE XYLASE ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S73809; AAB32873.1; -.
 DR HSP; P24300; 2XIS.
 DR PROSITE; PS00172; XYLASE_ISOMERASE_1; 1.
 DR PROSITE; PS00173; XYLASE_ISOMERASE_2; 1.
 DR PFAM; PF00259; Xylase_isom; 1.
 KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.
 FT ACT_SITE 54 54 BY SIMILARITY.
 FT METAL 181 181 MAGNESIUM (BY SIMILARITY).
 FT METAL 217 217 MAGNESIUM (BY SIMILARITY).
 FT METAL 245 245 MAGNESIUM (BY SIMILARITY).
 FT METAL 287 287 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 388 AA; 42675 MW; 5E05C770 CRC32;

Query Match 90.4%; Score 1834; DB 1; Length 388;
 Best Local Similarity 89.9%; Pred. NO. 1.6e-136;
 Matches 349; Conservative 16; Mismatches 21; Indels 2; Gaps 2;

QY 1 MNQPTPEDRFTFGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
 Db 1 MSIQPTPEDKFTFGLTWGVQGRDPFGDTRPALDPAESVRLAELGAYGVTFHDDDLIP 60
 QY 60 FGASDTEREAHVKFRQALDAGTMTVPMTATNLTHPVEKAGAFANDRAVRRYALRKTI 119

Db 61 FGATDSERAEHKKFRQGLDETGMKVPMTATNLTHPVEKAGAFANDRAVRRYALRKTI 120
 QY 120 RNIDLAVELGAKVYVWGGREGAAGKADYRAALDRKKEAFDILLGEYVTSQGYDIRFAI 179
 Db 121 RNIDLAVELGAKVYVWGGREGAAGKADYRAALDRKKEAFDILLGEYVTSQGYDIRFAI 180
 QY 180 EPKNEPRGDILLPTTGHAFIERLERPELYGVNPEVGHQMAGLNPHGIAQALWAGK 239
 Db 181 EPKNEPRGDILLPTTGHAFIERLERPELYGVNPEVGHQMAGLNPHGIAQALWAGK 240
 QY 240 LFHIDLNGSGIKYDODLRFGAGDLRAAFWLVLLSAGWEGPRHDFKPPRTEDIDGVM 299
 Db 241 LFHIDLNGSGIKYDODLRFGAGDLRAAFWLVLLSAGWEGPRHDFKPPRTEDIDGVM 300
 QY 300 ASAAGCMRNLYLILKRAAFAFRADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDF 359
 Db 301 ASAAGCMRNLYLILKRAAFAFRADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDF 360
 QY 360 DVD-AAARGMAFERLDQALDHLGARG 386
 Db 361 DPDAARGMAFERLDQALDHLGARG 388

RESULT 7
 XYL_A_STRRO STANDARD; PRT; 393 AA.
 AC P22857;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE XYLASE ISOMERASE (EC 5.3.1.5).
 GN XYL.
 OS Streptomyces rochei (Streptomyces parvullus).
 OC Bacteria; Firmicutes; Actinobacteri; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S-41;
 RX MEDLINE; 91136786.
 RA KIKUCHI T., ITOH Y., KASUMI T., FUKAZAWA C.;
 RT "Molecular cloning of the xylA gene encoding xylose isomerase from Streptomyces griseofuscus S-41: primary structure of the gene and its product".
 RL Agric. Biol. Chem. 54:2469-2472(1990).
 CC -!- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
 CC -!- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.
 CC -!- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SIMILARITY: BELONGS TO THE XYLASE ISOMERASE FAMILY.
 DR PIR; JN0086; ISSMXR.
 DR HSP; P37031; IDXI.
 DR PROSITE; PS00172; XYLASE_ISOMERASE_1; 1.
 DR PROSITE; PS00173; XYLASE_ISOMERASE_2; 1.
 DR PFAM; PF00259; Xylase_isom; 1.
 KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT ACT_SITE 53 53 MAGNESIUM (BY SIMILARITY).
 FT METAL 179 179 MAGNESIUM (BY SIMILARITY).
 FT METAL 215 215 MAGNESIUM (BY SIMILARITY).
 FT METAL 243 243 MAGNESIUM (BY SIMILARITY).
 FT METAL 284 284 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 393 AA; 43513 MW; 941FD6DE CRC32;

Query Match 75.6%; Score 1534.5; DB 1; Length 393;
 Best Local Similarity 79.2%; Pred. NO. 4.6e-113;
 Matches 305; Conservative 21; Mismatches 24; Indels 35; Gaps 8;

QY 2 NYQPTPEDRFTFGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 60
 Db 1 SFQPTPEDKFTFGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 60

QY 61 GASDTEREAKVHKFRQALDQATGVTVPMTATNLTHPVPFKAGATANDRAVRRYALRKRTIR 120
 DB 61 GSSDTERESHKFRQALDQATGVTVPMTATNLTHPVPFK-DRETANDRDVRAVAVRKRTIR 119
 QY 121 NIDLAELGAKTVVANGREGSAGGAKOVRAALDRKKEAFDILGGEVTSOGYDIRFAIE 180
 DB 120 NIDLAELGAKTVVANGREGSAGGAKOVRAALDRKKEAFDILGGEVTSOGYDIRFAIE 179
 QY 181 PKPNEPRGDILLPTIGHALAFIERLERPELYGNPEVGHQEMAGLNFPHGIAQALWAGKL 240
 DB 180 PKPNEPRGDILLPTIGHALAFIERLERPELYGNPEVGHQEMAGLNFPHGIAQALWAGKL 239
 QY 241 FHDILNGSQSIKYDQDL---RFGAGDLRAFWLVDLLESAGWEGPRHDFKPPRTEDIDG 297
 DB 240 FHDILNGSQSIKYDQCGRRRPPAGV---LVVDLLESAGWEGPRHDFKPPRTEDIDG 295
 QY 298 VWASAGCMRNYLLIKE-----RAAFRADPEVQALRAARLDQALAEPTAAD 344
 DB 296 VWASAGCMRNYLLIKOPRPPAPTRCRRASA----PRVWTSWFSRPL-----AD 343
 QY 345 GLOALLADRTAYEDFDVD-AAARGM 368
 DB 344 GLEALLADRTAFEDFVEAAARGM 368

RESULT 8
 XYLA.ACTMI STANDARD; PRT; 393 AA.
 AC P12851;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE XYLOSE ISOMERASE (EC 5.3.1.5).
 GN XYLA OR XI.
 OS Actinoplanes missouriensis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Micromonosporineae; Micromonosporaceae;
 OC Actinoplanes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DSM 43046;
 RX MEDLINE; 90016811.
 RA AMORE R., HOLLENBERG C.P.;
 RT "xylose isomerase from Actinoplanes missouriensis: primary structure
 of the gene and the protein.";
 RL Nucleic Acids Res. 17:7515-7515(1989).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE; 89184498.
 RA REY F., JENKINS J., JANIN J., LASTERS I., ALARD P., CLAESSENS M.,
 RA MATTHYSSENS G., WODAK S.G.;
 RT "Structural analysis of the 2.8 A model of xylose isomerase from
 Actinoplanes missouriensis.";
 RL Proteins 4:165-172(1988).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE; 92304915.
 RA JENKINS J., JANIN J., REY F., CHIADMI M., VAN TILBEURGH H.,
 RA LASTERS I., DE MAEYER M., VAN BELLE D., WODAK S.J., LAUWEREYS M.,
 RA STANSENS P., MRABET N.T., SNAUWAERT J., MATTHYSSENS G.,
 RA LAMBEIR A.-M.;
 RT "Protein engineering of xylose (glucose) isomerase from Actinoplanes
 missouriensis. 1. Crystallography and site-directed mutagenesis of
 metal binding sites.";
 RL Biochemistry 31:5449-5458(1992).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (4.1 ANGSTROMS).
 RA RAMIN M., SHEPARD W., FOURME R., KAHN R.;
 RL Submitted (JUN-1998) to the PDB data bank.
 RN [5]
 RP MUTAGENESIS.
 RX MEDLINE; 92304916.

RA LAMBEIR A.-M., LAUWEREYS M., STANSENS P., MRABET N.T., SNAUWAERT J.,
 RA VAN TILBEURGH H., MATTHYSSENS G., LASTERS I., DE MAEYER M.,
 RA WODAK S.J., JENKINS J., CHIADMI M., JANIN J.;
 RT "Protein engineering of xylose (glucose) isomerase from Actinoplanes
 missouriensis. 2. Site-directed mutagenesis of the xylose binding
 site.";
 RL Biochemistry 31:5459-5466(1992).
 RN [6]
 RP MUTAGENESIS.
 RX MEDLINE; 92304917.
 RA VAN TILBEURGH H., JENKINS J., CHIADMI M., JANIN J., WODAK S.J.,
 RA MRABET N.T., LAMBEIR A.-M.;
 RT "Protein engineering of xylose (glucose) isomerase from Actinoplanes
 missouriensis. 3. Changing metal specificity and the pH profile by
 site-directed mutagenesis.";
 RL Biochemistry 31:5467-5472(1992).
 CC -!- CATALYTIC ACTIVITY: D-XYLOSE = D-XYLULOSE.
 CC -!- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X16042; CAA34164.1;
 DR EMBL; AL0241; CAA00884.1;
 DR PIR; S05998; ISMAXM.
 DR PDB; 1XIN; 15-JUL-93.
 DR PDB; 2XIN; 15-JUL-93.
 DR PDB; 3XIN; 15-JUL-93.
 DR PDB; 5XIN; 15-JUL-93.
 DR PDB; 1XIM; 15-APR-93.
 DR PDB; 2XIM; 15-APR-93.
 DR PDB; 3XIM; 15-APR-93.
 DR PDB; 4XIM; 15-JUL-93.
 DR PDB; 5XIM; 15-JUL-93.
 DR PDB; 6XIM; 15-JUL-93.
 DR PDB; 7XIM; 15-JUL-93.
 DR PDB; 8XIM; 15-JUL-93.
 DR PDB; 9XIM; 15-JUL-93.
 DR PDB; 1BHW; 18-NOV-98.
 DR PROSITE; PS00172; XYLOSE ISOMERASE_1; 1.
 DR PROSITE; PS00173; XYLOSE ISOMERASE_2; 1.
 DR PFW; PF00259; Xylose_isom; 1.
 KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.
 FT INIT MET 0 0
 FT ACT SITE 53 53
 FT METAL 180 180
 FT METAL 216 216
 FT METAL 244 244
 FT METAL 291 291
 FT HELIX 6 8
 FT STRAND 10 13
 FT HELIX 14 17
 FT TURN 18 18
 FT STRAND 23 23
 FT TURN 24 25
 FT STRAND 26 26
 FT HELIX 35 45
 FT TURN 46 46
 FT STRAND 49 51
 FT STRAND 53 53
 FT HELIX 54 57
 FT TURN 60 61
 FT HELIX 64 81
 FT TURN 82 82
 FT STRAND 84 84

20

```

FT TURN 279 280
FT STRAND 289 291
FT TURN 297 298
FT TURN 301 327
FT HELIX 301 327
FT TURN 329 337
FT TURN 338 341
FT HELIX 342 344
FT TURN 350 351
FT HELIX 354 359
FT TURN 361 366
FT HELIX 369 372
FT TURN 373 374
FT HELIX 379 390
FT TURN 391 392
SQ SEQUENCE 394 AA: 43160 MW; BCB86DC2 CRC32;

Query Match 62.9%; Score 1277; DB 1; Length 394;
Best Local Similarity 62.5%; Pred. No. 6.6e-93;
Matches 245; Conservative 52; Mismatches 85; Indels 10; Gaps 4;

QY 4 QPTPEDRTFTGLTVMGWRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIPFGA 62
DB 3 QPTPADHFTFTGLTVMGWRDPFGDTRPALDPVD-VAHVHKLAEGLGAYGVTFHDDDLIPFGA 62

QY 63 SDTEREAHVKFRQALDQATGTMVPMATNLTFTHPVFKAGAFANDRAVRRYALRKRTIRNI 122
DB 63 TEAREKILGDFENQALKDTGLKVPMTNLTSHPVFKDGGFTSNDRSIRRFALAKVLHNI 122

QY 123 DLAVELGAKVYVANGREGAESGAADVRAALDRMKEAFDGLLGEYVTSQGYDIRFAIEPK 182
DB 123 DLAAEMGAETFTVMGREGSEYDGSKDLAALDRMREGVDTAAGYIKDKGVNLRALIEPK 182

QY 183 PNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGKLFH 242
DB 183 PNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGKLFH 242

QY 243 IDLNQSGIKYDQDLRFAGADLRAAFWLVLDLES- - - - -GWEGPRHFDKPPRTEDID 296
DB 243 IDLNQSGIKYDQDLRFAGADLRAAFWLVLDLES- - - - -GWEGPRHFDKPPRTEDID 296

QY 297 GYWSAAGCMRNYLILKERAARAFADPEVQEAALRAARDQLAEPT- - -AADGLQALLADRT 354
DB 303 GYWSAKANMSHYLLKERAARAFADPEVQEAALRAARDQLAEPT- - -AADGLQALLADRT 354

QY 355 AYEDFDVDAAR-RGNAPERLDQLANDHLLGAR 385
DB 363 SFAGFDAAEAARNAFIRLNQLAIEHLIGSR 394

RESULT 11
XYLA_THETH STANDARD; PRT; 387 AA.
AC P26997;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91216978.
RA DEKKER K., YAMAGATA H., SAKAGUCHI K., UDAKA S.;
RT "Xylose (glucose) isomerase gene from the thermophile Thermus
RT thermophilus: cloning, sequencing, and comparison with other
RT thermostable xylose isomerases."
RL J. Bacteriol. 173:3078-3083(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RC STRAIN=HB8;
RX MEDLINE; 99264393.

CHANG C., PARK B.C., LEE D.-S., SUH S.W.;
"Crystal structures of thermostable xylose isomerases from Thermus
calidophilus and Thermus thermophilus: possible structural
determinants of thermostability.";
J. Mol. Biol. 288:623-634(1999).
CC -1- CATALYTIC ACTIVITY: D-XYLOSE = D-XYLULOSE.
CC -1- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D90256; BAA14301.1; -.
DR PIR; A39404; A39404.
DR PDB; 1BX8; 16-FEB-99.
DR PROSITE; PS00174; XYLOSE_ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; 1.
DR PFAM; PF00259; Xylose_isom; 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.
FT ACT_SITE 53 53 BY SIMILARITY.
FT METAL 180 180 MAGNESIUM (BY SIMILARITY).
FT METAL 216 216 MAGNESIUM (BY SIMILARITY).
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 286 286 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 387 AA; 43906 MW; B72CCEC7 CRC32;

Query Match 55.5%; Score 1127; DB 1; Length 387;
Best Local Similarity 58.5%; Pred. No. 3.6e-81;
Matches 228; Conservative 43; Mismatches 109; Indels 10; Gaps 3;

QY 3 YQPTPEDRTFTGLTVMGWRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIPFG 61
DB 2 YEPKPEHRFTFTGLTVMGWRDPFGDTRPALDPV-VYVHKLAEGLGAYGVNLDLIPRG 61

QY 62 ASDTEREAHVKFRQALDQATGTMVPMATNLTFTHPVFKAGAFANDRAVRRYALRKRTIRN 121
DB 62 TPQERQIVRFRKALDETGLKVPMTANLFSDFKAGAFSPDPWVRAVALRKSLST 121

QY 122 IDLAVELGAKVYVANGREGAESGAADVRAALDRMKEAFDGLLGEYVTSQGYDIRFAIEP 181
DB 122 MDLGAELGAETFTVMGREGAEVATGKARKVMDVWRVREALNFMAYAEADQGYRFALEP 181

QY 182 KPNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGKLF 241
DB 182 KPNEPRGDIYFATVGMALFHTLDRPERFGLNPEFHEHETMAGLNFMHVAQAADAGKLF 241

QY 242 HIDLNQSGIKYDQDLRFAGADLRAAFWLVLDLESAGWEGPRHFDKPPRTEDIDGYNAS 301
DB 242 HIDLNQSRMSRFDDDLRFGSENKAAFFVLDLLESSGQGRHFDHALRTEDEGWWAF 301

QY 302 AAGCMRNYLILKERAARAFADPEVQEAALRAARDQLAEPT- - -LAETPAADGLQALLADRTAY 356
DB 302 ARGCMRTYLLKERAARAFADPEVQEAALRAARDQLAEPT- - -LAETPAADGLQALLADRTAY 356

QY 357 EDFDVAARAGMAFERLDQLANDHLLGARG 386
DB 362 E- - -AKRRRGYALERLDQLAVEYLLGVRG 387

RESULT 12
XYLA_THETH STANDARD; PRT; 387 AA.
AC P56681;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)

```

RESULT 13

```

Query Match      16.2%; Score 329; DB 1; Length 448;
Best Local Similarity 29.5%; Pred. NO. 1.3e-18;
Matches 104; Conservative 55; Mismatches 160; Indels 34; Gaps

Qy 10 RFTGLW-TVCGQRDPFGDAT-----RPALDPVDVQR-----LAELGAVGYTFHD 54
|| || || || || || || || || || || || || || || || || || || || ||
Db 41 RFAYVHTFDADGKDPFGDGTFRANRRTHTPLDKAKARAEEFEELGVPYFCFHD 100
|| || || || || || || || || || || || || || || || || || || || ||
Qy 55 DDLIPFGASDTERAHYKRV----FRQALDAGTVPVMTNLTTHPVFKAGATANDRAW 110
|| || || || || || || || || || || || || || || || || || || || ||

```

```

Query Match      16.0%; Score 324.5; DB 1; Length 435;
Best Local Similarity 30.4%; Pred. No. 2.8e-18;
Matches 109; Conservative 48; Mismatches 162; Indels 39; Gaps
Qy 10 RFEFLG-TVGWGRDPFGDAT-----RPAIDPVDV--QRLEIAGYGVTFHDD 55
      |||
      |||
Db 42 RFSVAFHHTTDEGSDPFGDQTYORNWEGSTPMTAKNRVDATFEIIEKLGJFYCFCHDV 103
      |||
      |||

```

Search completed: March 22, 2000, 23:46:36
Job time: 187 sec

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1892.5	93.3	389	1	ISSXV	xylose isomerase
2	1877	92.5	388	1	B41339	xylose isomerase
3	1872.5	92.3	388	1	S28986	xylose isomerase
4	1831	90.2	388	2	JC1031	xylose isomerase
5	1539.5	75.9	394	1	ISWXR	xylose isomerase
6	1353.5	66.7	394	1	ISWAXA	xylose isomerase
7	1351.5	66.6	394	1	ISWAXA	xylose isomerase
8	1282	63.2	395	1	S16214	xylose isomerase
9	1127	55.5	387	2	A39404	xylose isomerase
10	333.5	16.4	449	2	S18561	xylose isomerase
11	318	15.7	439	2	I40806	xylose isomerase
12	317	15.6	444	2	A72225	xylose isomerase
13	314	15.5	449	2	JC1137	xylose isomerase
14	312.5	15.4	439	2	S16530	xylose isomerase
15	310.5	15.3	438	1	ISGLXM	xylose isomerase
16	301.5	14.9	438	1	ISGLXH	xylose isomerase
17	297.5	14.7	445	1	ISBSXS	xylose isomerase
18	277.5	13.7	440	1	ISRCX1	xylose isomerase
19	273.5	13.5	440	1	ISRBX	xylose isomerase
20	239.5	11.8	439	1	ISHLX	xylose isomerase
21	235	11.6	479	2	S65466	xylose isomerase
22	194	9.6	34	2	PS0335	xylose isomerase
23	179	8.8	34	2	PS0336	xylose isomerase
24	138.5	6.8	383	2	A72299	xylose isomerase
25	107	5.3	233	2	H69130	sugar isomerase - conserved hypothetical
26	98.5	4.9	1061	2	S37667	trac-1 protein - E
27	97	4.8	1381	1	Q0B845	major capsid protein
28	96	4.7	415	1	S75766	hypothetical protein
29	96	4.7	800	2	A29003	cellulase (EC 3.2.1)
30	95.5	4.7	367	2	S65502	potassium channel

Db 241 LFHIDLNGSGIKYDQDLRFAGDRLRAAFWLVLDLESAGYEGPRHFDKPPRTDIDGVW 300
Qy 300 ASAGCMRNILKERAFAFRADPEVQALRAALDQAEPTAADGLQALLADRTAYEDF 359
Db 301 ASAGCMRNILKERAFAFRADPEVQALRAALDQAEPTAADGLQALLADRTAYEDF 360
Qy 360 DVDAARAGMA--FERLDQALMDHLLGARG 386
Db 361 DVDAARAGMA--FERLDQALMDHLLGARG 389
RESULT 2
B41339
xylose isomerase (EC 5.3.1.5) - Streptomyces rubiginosus
C:Species: Streptomyces rubiginosus
C:Date: 28-May-1992 #sequence_revision 06-Dec-1996 #text_change 18-Jun-1999
C:Accession: B41339
R:Wong, H.C.; Ting, Y.; Lin, H.C.; Reichert, F.; Myambo, K.; Watt, K.W.K.; Toy, P.L.; Du
J. Bacteriol. 173, 6849-6858, 1991
A:Title: Genetic organization and regulation of the xylose degradation genes in Streptom
A:Reference number: A41339; MUID:92041569
A:Accession: B41339
A:Molecule type: DNA
A:Residues: 1-388 <WON>
A:Cross-references: GB:M73789; NID:g153532; PIDN:AAA26838.1; PID:g153534
A:Note: parts of this sequence, including the amino and carboxyl ends of the mature prot
R:Carroll, H.L.; Glusker, J.P.
Submitted to the Brookhaven Protein Data Bank, October 1990
A:Reference number: A50791; PDB:7XIA
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 1-40, 'Q', 42-387
R:Carroll, H.L.; Glusker, J.P.
submitted to the Brookhaven Protein Data Bank, October 1990
A:Reference number: A50808; PDB:8XIA
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, with D-xylose, residues 1-
R:Carroll, H.L.; Glusker, J.P.; Burger, V.; Manire, F.; Tritsch, D.; Biellmann, J.F.
Proc. Natl. Acad. Sci. U.S.A. 86, 4440-4444, 1989
A:Title: X-ray analysis of D-xylose isomerase at 1.9 angstroms: native enzyme in complex
A:Reference number: A58422; MUID:89282788
A:Contents: annotation; X-ray crystallography, 1.9 angstroms
R:Whitlow, M.; Howard, A.J.
submitted to the Brookhaven Protein Data Bank, March 1991
A:Reference number: A50703; PDB:4XIS
A:Contents: annotation; X-ray crystallography, 1.60 angstroms, with xylose and manganese
R:Whitlow, M.; Howard, A.J.; Finzel, B.C.; Poulos, T.L.; Winborne, E.; Gilliland, G.L.
Proteins 9, 153-173, 1991
A:Title: A metal-mediated hydride shift mechanism for xylose isomerase based on the 1.6
A:Reference number: A58421; MUID:91172741
A:Contents: annotation; X-ray crystallography, 1.60 angstroms
C:Complex: homotetramer
C:Function:
A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopentose
A:Pathway: xylose metabolism
C:Superfamily: xylose isomerase
C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose me
F:1-388/Product: xylose isomerase #status experimental <MAT>
F:54,57/Active site: His, Asp #status experimental
F:90,183/Binding site: substrate (Thr, Lys) #status experimental
F:181,217,245,287/Binding site: magnesium (Glu, Glu, Asp, Asp) #status experimental
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status experimental

Query Match 92.5%; Score 1877; DB 1; Length 388;
Best Local Similarity 92.3%; Pred. No. 4.5e-136;
Matches 358; Conservative 14; Mismatches 14; Indels 2; Gaps 2;
Qy 1 MNQPTPEDRTFFGLTWGVGWRDPFGDTRPALDPVD-VORLAEAGYGVTFHDDDLIP 59
Db 1 MNQPTPEDRTFFGLTWGVGWRDPFGDTRPALDPVESVRLAEALGAGVTFHDDDLIP 60
Qy 60 FGASDTEREAHVKRFRQALDGMVPMATTNLFTHPVFKAGFTANDRAVRRYALRXTI 119
Db 61 FGSSDTEREAHVKRFRQALDGMVPMATTNLFTHPVFKAGFTANDRAVRRYALRXTI 120

Qy 120 RNIDLAVELGAKVYVAMGREGAESGAADKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
Db 121 RNIDLAVELGAKVYVAMGREGAESGAADKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 180
Qy 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
Db 181 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 240
Qy 240 LFHIDLNGSGIKYDQDLRFAGDRLRAAFWLVLDLESAGYEGPRHFDKPPRTDIDGVW 299
Db 241 LFHIDLNGSGIKYDQDLRFAGDRLRAAFWLVLDLESAGYEGPRHFDKPPRTDIDGVW 300
Qy 300 ASAGCMRNILKERAFAFRADPEVQALRAALDQAEPTAADGLQALLADRTAYEDF 359
Db 301 ASAGCMRNILKERAFAFRADPEVQALRAALDQAEPTAADGLQALLADRTAYEDF 360
Qy 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
Db 361 DVDAARAGMAFERLDQALMDHLLGARG 388
RESULT 3
S28986
xylose isomerase (EC 5.3.1.5) - Streptomyces olivochromogenes
C:Species: Streptomyces olivochromogenes
C:Date: 25-Feb-1994 #sequence_revision 06-Dec-1996 #text_change 10-Sep-1999
C:Accession: S28986
R:Farber, G.K.; Glasfeld, A.; Tiraby, G.; Ringe, D.; Petsko, G.A.
Biochemistry 28, 7289-7297, 1989
A:Title: Crystallographic studies of the mechanism of xylose isomerase.
A:Reference number: S28986; MUID:90057422
A:Contents: X-ray crystallography, 3.0 angstroms
A:Accession: S28986
A:Molecule type: protein; DNA
A:Residues: 1-389 <FAR>
R:Farber, G.; Petsko, G.
submitted to the Brookhaven Protein Data Bank, February 1989
A:Reference number: A50637; PDB:3XIA
A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 2-378
C:Complex: homodimer
C:Function:
A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopento
A:Pathway: xylose metabolism
C:Superfamily: xylose isomerase
C:Keywords: homodimer; intramolecular oxidoreductase; isomerase; magnesium; xylose me
F:2-389/Product: xylose isomerase #status experimental <MAT>
F:54,57/Active site: His, Asp #status experimental
F:90,183/Binding site: substrate (Thr, Lys) #status experimental
F:181,217,245,287/Binding site: magnesium (Glu, Glu, Asp, Asp) #status experimental
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status experimental

Query Match 92.3%; Score 1872.5; DB 1; Length 389;
Best Local Similarity 91.5%; Pred. No. 1e-135;
Matches 356; Conservative 15; Mismatches 15; Indels 3; Gaps 2;
Qy 1 MNQPTPEDRTFFGLTWGVGWRDPFGDTRPALDPVD-VORLAEAGYGVTFHDDDLIP 59
Db 1 MSYQTPEDRTFFGLTWGVGWRDPFGDTRPALDPVDVORLAEAGYGVTFHDDDLIP 60
Qy 60 FGASDTEREAHVKRFRQALDGMVPMATTNLFTHPVFKAGFTANDRAVRRYALRXTI 119
Db 61 FGSSDTEREAHVKRFRQALDGMVPMATTNLFTHPVFKAGFTANDRAVRRYALRXTI 120
Qy 120 RNIDLAVELGAKVYVAMGREGAESGAADKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
Db 121 RNIDLAVELGAKVYVAMGREGAESGAADKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 180
Qy 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
Db 181 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 240
Qy 240 LFHIDLNGSGIKYDQDLRFAGDRLRAAFWLVLDLESAGYEGPRHFDKPPRTDIDGVW 299

Db 241 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGYEGPRHLDFKPRTEIDIGVW 300
 QY 300 ASAGCMRNLYLKERAAAFRADEVEALRAARLDOLAEPDADGQALLADRTAYEDF 359
 Db 301 ASAGCMRNLYLKERAAAFRADEVEALRAARLDOLAEPDADGQALLADRTAYEDF 360
 QY 360 DVDAARGMA--FERLDOLAMHLLGARG 386
 Db 361 DVDAARGMAWFERLDOLAMHLLGARG 389

RESULT 4
 JC1031
 xylose isomerase (EC 5.3.1.5) - Streptomyces diastaticus
 C:Species: Streptomyces diastaticus
 C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 06-Dec-1996
 C:Accession: JC1031
 R:Wang, Y.Z.; Huang, Z.; Dai, X.H.; Liu, J.; Cui, T.; Niu, L.W.; Wang, C.; Xu, X.
 Chinese J. Biotechnol. 10, 118-123, 1994
 A:Title: DNA sequence of xylose isomerase gene from Streptomyces diastaticus No.7 strain
 A:Reference number: JC1031
 A:Contents: No.7, M1033
 A:Accession: JC1031
 A:Molecule type: DNA
 A:Residues: 1-388 <WAN>
 C:Superfamily: xylose isomerase
 C:Keywords: Intramolecular oxidoreductase; isomerase

Query Match 90.2%; Score 1831; DB 2; Length 388;
 Best Local Similarity 89.7%; Pred. No. 1.5e-132;
 Matches 348; Conservative 17; Mismatches 21; Indels 2; Gaps 2;
 QY 1 MNYQTPEDRFTGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTHDDDLIP 59
 Db 1 MSQTPEDRFTGLTWVGWGRDPFGDTRPALDPVDVORLAELGAYGVTHDDDLIP 60
 QY 60 FGASDTEREAHVRFKROALDGMTVPMATNLTFTHPVKAGAFANDRAVRYALRKT 119
 Db 61 FGATDSERAEHFRKROALDGMTVPMATNLTFTHPVKAGAFANDRAVRYALRKT 120
 QY 120 RNTDLAVELGAKYVAVWGREGAESGAADKVRALDRMKEAFDLGLGYVTSQGYDIRFAI 179
 Db 121 RNTDLAVELGAKYVAVWGREGAESGAADKVRALDRMKEAFDLGLGYVTSQGYDIRFAI 180
 QY 180 EKPNEPRGDIILLPTIGHALAFIERLPERLYGVNPEVGEQWAGLNFPHGIAQALWAGK 239
 Db 181 EKPNEPRGDIILLPTIGHALAFIERLPERLYGVNPEVGEQWAGLNFPHGIAQALWAGK 240
 QY 240 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGYEGPRHDFKPRTEIDIGVW 299
 Db 241 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGYEGPRHDFKPRTEIDIGVW 300
 QY 300 ASAGCMRNLYLKERAAAFRADEVEALRAARLDOLAEPDADGQALLADRTAYEDF 359
 Db 301 ASAGCMRNLYLKERAAAFRADEVEALRAARLDOLAEPDADGQALLADRTAYEDF 360
 QY 360 DVD-AAARGWAFERLDOLAMHLLGARG 386
 Db 361 DPDAARGWAFERLDOLAMHLLGARG 388

RESULT 5
 ISSMR
 xylose isomerase (EC 5.3.1.5) - Streptomyces rochei
 N:Alternate names: xylose ketol-isomerase
 C:Species: Streptomyces rochei
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 07-May-1999
 C:Accession: JN0086
 R:Kikuchi, T.; Itoh, Y.; Kasumi, T.; Fukazawa, C.
 Agric. Biol. Chem. 54, 2469-2472, 1990
 A:Title: Molecular cloning of the xylA gene encoding xylose isomerase from Streptomyces

A:Reference number: JN0086; MUID:91136786
 A:Accession: JN0086
 A:Molecule type: DNA
 A:Residues: 1-394 <KIK>
 C:Comment: Phe-26, Phe-94, and Trp-136 have been demonstrated to be involved in facII
 C:Genetics:
 A:Gene: xylA
 A:Complex: homotetramer
 C:Function:
 A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopento
 A:Pathway: xylose metabolism
 C:Superfamily: xylose isomerase
 C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose
 F:2-394/Product: xylose isomerase #status predicted <MAT>
 F:54/57/Active site: His, Asp #status predicted
 F:90/182/Binding site: substrate (Thr, Lys) #status predicted
 F:180/216/244/285/Binding site: magnesium (Glu, Glu, Asp, Asp) #status predicted
 F:216/219/254/256/Binding site: magnesium (Glu, His, Asp, Asp) #status predicted
 Query Match 75.9%; Score 1539.5; DB 1; Length 394;
 Best Local Similarity 79.3%; Pred. No. 2.7e-110;
 Matches 306; Conservative 21; Mismatches 24; Indels 35; Gaps 8;
 QY 1 MNYQTPEDRFTGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTHDDDLIP 59
 Db 1 MSQTPEDRFTGLTWVGWGRDPFGDTRPALDPVDVORLAELGAYGVTHDDDLIP 60
 QY 60 FGASDTEREAHVRFKROALDGMTVPMATNLTFTHPVKAGAFANDRAVRYALRKT 119
 Db 61 FGSDTERESHARFROALDGMTVPMATNLTFTHPVKAFDRFTANDRAVRYALRKT 119
 QY 120 RNTDLAVELGAKYVAVWGREGAESGAADKVRALDRMKEAFDLGLGYVTSQGYDIRFAI 179
 Db 120 RNTDLAVELGAKYVAVWGREGAESGAADKVRALDRMKEAFDLGLGYVTSQGYDIRFAI 179
 QY 180 EKPNEPRGDIILLPTIGHALAFIERLPERLYGVNPEVGEQWAGLNFPHGIAQALWAGK 239
 Db 180 EKPNEPRGDIILLPTIGHALAFIERLPERLYGVNPEVGEQWAGLNFPHGIAQALWAGK 239
 QY 240 LFHIDLNGQSGIKYDQDL--REFAGDLRAAFWLVLDLLESAGYEGPRHDFKPRTEID 296
 Db 240 LFHIDLNGQSGIKYDQCGSRPPAGGV----LVVDLLESAGYEGPRHDFKPRTEID 295
 QY 297 GVNASAGCMRNLYLKE-----RAAFRADEVEALRAARLDOLAEPDAA 343
 Db 296 GVNASAGCMRNLYLKE-----RAAFRADEVEALRAARLDOLAEPDAA 343
 QY 344 DGLQALLADRTAYEDFVDV-AAARGM 368
 Db 344 DGLQALLADRTAYEDFVDV-AAARGM 369
 RESULT 6
 ISMAXA
 xylose isomerase (EC 5.3.1.5) - Ampullariella sp.
 N:Alternate names: xylose ketol-isomerase
 C:Species: Ampullariella sp.
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 06-Dec-1996
 C:Accession: A27756
 R:Saari, G.C.; Kumar, A.A.; Kawasaki, G.H.; Insley, M.Y.; O'Hara, P.J.
 J. Bacteriol. 169, 612-618, 1987
 A:Title: Sequence of the Ampullariella sp. strain 3876 gene coding for xylose isomera
 A:Reference number: A27756; MUID:87109047
 A:Accession: A27756
 A:Molecule type: DNA
 A:Residues: 1-394 <SAA>
 A:Cross-references: GB:M15050
 A:Experimental source: strain 3876
 C:Genetics:
 A:Start codon: GTG
 C:Complex: homotetramer
 C:Function:

A;Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopentose
A;Pathway: xylose metabolism
C;Superfamily: xylose isomerase
C;Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose me
F:54,57/Active site: His, Asp #status predicted
F:90,183/Binding site: substrate (Thr, Lys) #status predicted
F:181,217,245,292/Binding site: magnesium (Glu, Asp, Asp) #status predicted
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status predicted

Query Match 66.7%; Score 1353.5; DB 1; Length 394;
Best Local Similarity 66.0%; Pred. No. 4.3e-96;
Matches 260; Conservative 46; Mismatches 79; Indels 9; Gaps 4;

Qy 1 MNYQPTPEDRTFFGLTWVGWGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
Db 1 MSLQATPDOKFSGFLWTVGWAQDAFGDTRPVLPDIEAVHKLAEIGAYGVTFHDDDLVP 60

Qy 60 FGASDTEREAHVYKFRQALDGMTVPMATTNLFTHPVFKAGAFANDRAVRYALRXTI 119
Db 61 FGADAATRDGIVAGFSKALDETGLIVPMVTTNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120

Qy 120 RNIDLAELGAKVYVWGWREGAESAADKDYRAALDRMKEAFDILGEYVTSOGYDIRFAI 179
Db 121 ROMDLGAELGAKTLLVWGREGAEDYSAKDYGAAALDRYREALNLLAQYSEDQGYGLPFAI 180

Qy 180 EPKPNPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMGALNFPHGIAQALWAGK 239
Db 181 EPKPNPRGDIILLPTAGHAIQVQELERPELFGINPETHGQMSNLNFTQGIQAALWHKK 240

Qy 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLLESA-----GWGPRHDFKPPRTED 294
Db 241 LFHIDLNGHGPKEQDQDLVFGHGLLNAPSLVDLLENGPDGPGAYDPGRHFDYKPSRTED 300

Qy 295 IDGVWASAGCNRNLYILKERAAFRADPEVQEAALRAALDQAEPTAAG--LQALLAD 352
Db 301 FDGVWESAKDNTMYLLKRAKAFRADEPQVQAALAESKVDLPTPTLMPGETYADLLAD 360

Qy 353 RTAYEDFDVDA-AARGMAFERLDQLAMDLHLLGAR 385
Db 361 RSAPEDYDADAVGAKGFGVKNLQNLAIHLLGAR 394

RESULT 7
ISMAXH
xylose isomerase (EC 5.3.1.5) - Actinoplanes missouriensis
N;Alternate names: xylose ketol-isomerase
C;Species: Actinoplanes missouriensis
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
C;Accession: S05998
R;Amore, R.; Hollenberg, C.P.
Nucleic Acids Res. 17, 7515, 1989
A;Title: Xylose isomerase from Actinoplanes missouriensis: primary structure of the gene
A;Reference number: S05998; MUID:90016811
A;Accession: S05998
A;Molecule type: DNA
A;Residues: 1-394 <C&O>
A;Cross-references: EMBL:X16042; NID:G38870; PIDN:CAA34164.1; PID:G580713
A;Note: the authors translated the initiation codon GTG for residue 1 as Val
R;Janin, J.
submitted to the Brookhaven Protein Data Bank, April 1992
A;Reference number: A51549; PDB:6XIM
C;Contents: annotation; X-ray crystallography, 2.5 angstroms, with xylose and magnesium,
R;Rev, F.; Jenkins, J.; Janin, J.; Lasters, I.; Alard, P.; Claessens, M.; Matthysens, G
Proteins 4, 165-172, 1988
A;Title: Structural analysis of the 2.8 angstroms model of xylose isomerase from Actinop
A;Reference number: A58417; MUID:89184498
C;Contents: annotation; X-ray crystallography, 2.8 angstroms
C;Genetics:
A;Start codon: GTG
C;Function: homotetramer
A;Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopentose

A;Pathway: xylose metabolism
C;Superfamily: xylose isomerase
C;Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose
F:54,57/Active site: His, Asp #status experimental
F:90,183/Binding site: substrate (Thr, Lys) #status experimental
F:181,217,245,292/Binding site: magnesium (Glu, Asp, Asp) #status experimental
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status experimental

Query Match 66.6%; Score 1351.5; DB 1; Length 394;
Best Local Similarity 65.5%; Pred. No. 6.1e-96;
Matches 258; Conservative 51; Mismatches 76; Indels 9; Gaps 4;

Qy 1 MNYQPTPEDRTFFGLTWVGWGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
Db 1 MSVQATREDKESFGLWTVGWAQDAFGDTRTALDPVAVHKLAEIGAYGVTFHDDDLVP 60

Qy 60 FGASDTEREAHVYKFRQALDGMTVPMATTNLFTHPVFKAGAFANDRAVRYALRXTI 119
Db 61 FGSDAQTDRGIIAGFKALDETGLIVPMVTTNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120

Qy 120 RNIDLAELGAKVYVWGWREGAESAADKDYRAALDRMKEAFDILGEYVTSOGYDIRFAI 179
Db 121 ROMDLGAELGAKTLLVWGREGAEDYSAKDYSALDRYREALNLLAQYSEDQGYGLRFAI 180

Qy 180 EPKPNPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMGALNFPHGIAQALWAGK 239
Db 181 EPKPNPRGDIILLPTAGHAIQVQELERPELFGINPETHGQMSNLNFTQGIQAALWHKK 240

Qy 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLLESA-----GWGPRHDFKPPRTED 294
Db 241 LFHIDLNGHGPKEQDQDLVFGHGLLNAPSLVDLLENGPDGPGAYDPGRHFDYKPSRTED 300

Qy 295 IDGVWASAGCNRNLYILKERAAFRADPEVQEAALRAALDQAEPT--AADGLQALLAD 352
Db 301 YDGVWESAKDNTMYLLKRAKAFRADEPQVQAALAESKVDLPTPTLMPGEGYAEILLAD 360

Qy 353 RTAYEDFDVDA-AARGMAFERLDQLAMDLHLLGAR 385
Db 361 RSAPEDYDADAVGAKGFGVKNLQNLAIHLLGAR 394

RESULT 8
S16214
xylose isomerase (EC 5.3.1.5) - Arthrobacter sp.
N;Alternate names: D-glucose isomerase
C;Species: Arthrobacter sp.
C;Date: 21-Nov-1993 #sequence_revision 06-Dec-1996 #text_change 18-Jun-1999
C;Accession: S16214; S16212; S28154
R;Loviny-Anderson, T.; Shaw, P.C.; Shin, M.K.; Hartley, B.S.
Biochem. J. 277, 263-271, 1991
A;Title: D-Xylose (D-glucose) isomerase from Arthrobacter strain N.R.R.L. B3728. Gene
A;Reference number: S16213; MUID:91307531
A;Accession: S16214
A;Molecule type: DNA
A;Residues: 1-395 <BIO>
A;Cross-references: EMBL:X59466; NID:G39058; PIDN:CAA42073.1; PID:G39059
A;Experimental source: strain NRRL B3728
R;Smith, C.A.; Rangarajan, M.; Hartley, B.S.
Biochem. J. 277, 255-261, 1991
A;Title: D-Xylose (D-glucose) isomerase from Arthrobacter strain N.R.R.L. B3728. Puri
A;Reference number: S16212; MUID:91307530
A;Accession: S16212
A;Molecule type: protein
A;Residues: 2-21 <SMI>
A;Experimental source: strain NRRL B3728
A;Note: other peptide sequences were also determined
R;Siddiqui, K.S.; Rangarajan, M.; Hartley, B.S.; Kitmitto, A.; Panico, M.; Blench, I.
Biochem. J. 289, 201-208, 1993
A;Title: Arthrobacter D-xylose isomerase: partial proteolysis with thermolysin.
A;Reference number: S28154; MUID:93143669
A;Accession: S28154
A;Molecule type: protein

A:Residues: 337-348 <STD>
R:Henrick, K.; Collier, C.A.; Blow, D.M.
submitted to the Brookhaven Protein Data Bank, July 1989
A:Reference number: A50702; PDB:4XIA
A:Contents: annotation: X-ray crystallography, 2.3 angstroms, with D-sorbitol, residues
R:Henrick, K.; Collier, C.A.; Blow, D.M.
J. Mol. Biol. 208, 129-157, 1989
A:Title: Structures of D-xylose isomerase from *Arthrobacter* strain B3728 containing the
A:Reference number: A58418; MUID:89362453
A:Contents: annotation: X-ray crystallography, 2.3 angstroms
R:Collier, C.A.; Henrick, K.; Blow, D.M.
submitted to the Brookhaven Protein Data Bank, October 1991
A:Reference number: A51428; PDB:1XLC
A:Contents: annotation: X-ray crystallography, 2.5 angstroms, with D-xylose/xyulose and
R:Collier, C.A.; Henrick, K.; Blow, D.M.
J. Mol. Biol. 212, 211-235, 1990
A:Title: Mechanism for aldose-ketose interconversion by D-xylose isomerase involving ring
A:Reference number: A58419; MUID:90704542
A:Contents: annotation: X-ray crystallography, 2.5 angstroms
C:Genetics:
A:Gene: xylA
C:Complex: homotetramer
C:Function:
A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopentose
A:Pathway: xylose metabolism
C:Superfamily: xylose isomerase
C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose me
F:2-395/Product: xylose isomerase #status experimental <MAT>
F:54.57/Active site: His, Asp #status experimental
F:90.183/Binding site: substrate (Thr, Lys) #status experimental
F:181,217,245,293/Binding site: magnesium (Glu, Glu, Asp, Asp) #status experimental
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status experimental

Query Match 63.2%; Score 1282; DB 1; Length 395;
Best Local Similarity 62.3%; Pred. No. 1.2e-90;
Matches 246; Conservative 53; Mismatches 86; Indels 10; Gaps 4;

QY 1 MNYQPTPEDRFTFGLTWVGWGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
DB 1 MSVQPTPDHFTFGLTWVGWGTADPGVATRKNLDPVEAVHKLAEALGAYGITHDNDLIP 60

QY 60 FGASDTEREAHVKRFRQALDAGTMTVPMTTNLFTHPVKAGAFANDRAVRYALRKTI 119
DB 1 FQATEAREKILGDFNQALDKDTGLKVPMTTNLFSHPVFKDGGFTSNDRSIRFALAKVL 120

QY 120 RNIDLAELGAKYVAVWGREGAESGAADVRAALDRMKEAFDILLGEYVTSQGYDIREAI 179
DB 121 HNIDLAEMGAETFFVMWGRGSEYDGSGLAALDRMREGVDTAAGYTKDGYNLRIAL 180

QY 180 EPKPNPRGDIILPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGK 239
DB 181 EPKPNPRGDIILPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGK 240

QY 240 LFHIDLNGSGIKYDQDLRFAGAGDLRAAFWLVDLSEA-----GWEGPRHDFKPPRTE 293
DB 241 LFHIDLNGSGIKYDQDLRFAGAGDLRAAFWLVDLSEA-----GWEGPRHDFKPPRTE 300

QY 294 DIDGVWASAGCRNRYLLKERAFAFPDVEQALRAALDQALAEPT--AAGDLQALLA 351
DB 301 GYDGVWDSAKANNMRYLLKERAFAFPDVEQALRAALDQALAEPT--AAGDLQALLA 360

QY 352 DRTAVDFDVFDAAL-RGMAFERLDQALMDHLIGAR 385
DB 361 DNASFAGFAEAAAERNAFIRLNQALAEHLGSR 395

RESULT 9
A39404
xylose isomerase (EC 5.3.1.5) - *Thermus aquaticus*
N:Alternate names: xylose ketol-isomerase
C:Species: *Thermus aquaticus*
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993

C:Accession: A39404
R:Dekker, K.; Yamagata, H.; Sakaguchi, K.; Uda, S.
J. Bacteriol. 173, 3078-3083, 1991
A:Title: Xylose (glucose) isomerase gene from the thermophile *Thermus thermophilus*: c
A:Reference number: A39404; MUID:91216978
A:Accession: A39404
A:Molecule type: DNA
A:Residues: 1-387 <DEK>
C:Superfamily: xylose isomerase
C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; pentos

Query Match 55.5%; Score 1127; DB 2; Length 387;
Best Local Similarity 58.5%; Pred. No. 8.3e-79;
Matches 228; Conservative 43; Mismatches 109; Indels 10; Gaps 3;

QY 3 YQPTPEDRFTFGLTWVGWGRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIPFG 61
DB 2 YEPKPEHRTFGLTWVGWGRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIPFG 61

QY 62 ASDTEREAHVKRFRQALDAGTMTVPMTTNLFTHPVKAGAFANDRAVRYALRKTI 121
DB 62 TPOQERDQIVRRFKALDETGLKVPMTTNLFTSDPAFKDGAFTSPDPWVRAVALRKSL 121

QY 122 IDLAELGAKYVAVWGREGAESGAADVRAALDRMKEAFDILLGEYVTSQGYDIREAI 181
DB 122 MDLGAELGAEIYVWVGREGAEVATGKARKVMDVWYREALNFMAYAEQGYGFAL 181

QY 182 KPNPRGDIILPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGK 241
DB 182 KPNPRGDIILPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGK 241

QY 242 HIDLNGSGIKYDQDLRFAGAGDLRAAFWLVDLSEAAGWEGPRHDFKPPRTEIDG 301
DB 242 HIDLNGSGIKYDQDLRFAGAGDLRAAFWLVDLSEAAGWEGPRHDFKPPRTEIDG 301

QY 302 AGCMRNRYLLKERAFAFPDVEQALRAALDQ-----LAEPTAAGDLQALLADRTAY 356
DB 302 ARGCMRTYLLKERAFAFPDVEQALRAALDQ-----LAEPTAAGDLQALLADRTAY 356

QY 357 EDFDVAAGMAFERLDQALMDHLIGAR 386
DB 362 E-----AKRRRGYALERLDQALVEYLLGVRG 387

RESULT 10
S18561
xylose isomerase (EC 5.3.1.5) - *Lactobacillus pentosus*
C:Species: *Lactobacillus pentosus*
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C:Accession: S18561
R:Lokman, B.C.; van Santen, P.; Verdoes, J.C.; Kruse, J.; Leer, R.J.; Posno, M.; Pou
Mol. Gen. Genet. 230, 161-169, 1991
A:Title: Organization and characterization of three genes involved in D-xylose catabo
A:Reference number: S18560; MUID:92079891
A:Accession: S18561
A:Molecule type: DNA
A:Residues: 1-449 <LOK>
A:Cross-references: EMBL:N57384; NID:g149604; PIDN:AAA25258.1; PID:g149606
C:Genetics:
A:Gene: xylA
C:Superfamily: xylose isomerase
C:Keywords: intramolecular oxidoreductase; isomerase; magnesium; xylose metabolism
F:103,106/Active site: His, Asp #status predicted
F:143,236/Binding site: substrate (Thr, Lys) #status predicted
F:234,270,298,342/Binding site: magnesium (Glu, Glu, Asp, Asp) #status predicted

Query Match 16.4%; Score 333.5; DB 2; Length 449;
Best Local Similarity 29.8%; Pred. No. 3.9e-18;
Matches 122; Conservative 61; Mismatches 150; Indels 77; Gaps 18;

Matches 101; Conservative 53; Mismatches 149; Indels 35; Gaps 11;

```

Qy 10 RFTFGLW-TVGMQGRDRPGDAT--RP---ALDPVDVOR-----LAELGAYGVTFHD 54
    ||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 43 RFSIAYHITFTADTDQFGKATMORPNHHTDPMIDIAKARVEAAFEFEDKINAFYCFHD 102
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 55 DDLIPFGASDTEREAH-----VKRFQALDATGMTVPMTATNLTHTPVFKAGAFANDR 108
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 103 RDIAPEG--DTLRETNKNLDTIVAMINDYLATSKTKVLWGTAFLFSNPRFVHGASTSCNA 160
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 109 AVRRYALRKTIRNIDLAVELGAKVYVAVGGREGAEGSAKDVRAALDRMKEAFDILLGEYV 168
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 161 DVFAYSAAQVKALEITKELGGENYVFWGGREGYETLNTDMEFELDNFAFLHMAVDYA 220
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 169 TSQGVDIRFAIEPKPNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEHQMAGLNFP 228
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 221 KEIGFEGQFLIEPKPKEPTKKHQYDFDVANVLAFLKKYDLDKFKVNIENHATLAFHDFQ 280
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 229 HGIAQALWAGKLFHIDLN-GOSGIKYDQDLRFAGDLEAAFLVLDLLESAGWE-GPRHFD 286
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 281 HELRYARINGVLGSIDANTGDMLLGWDTD-QPTDIRMTTLAMVEVIKMGGFDKGGLNFD 339
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 287 FKPPRT-----ED-----IDGVWASAAGCMRNLYILKER 315
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 340 AKVRRASPEDEFLGHIAGMDAFAGKFKVAYKLVKDR 377
    |: | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: March 22, 2000, 23:44:40
Job time: 776 sec